THE UNIVERSITY OF QUEENSLAND

Forecasting The Price Of Iron Ore

Student Name: James Omundsen

Course Code: MINE4123

Supervisor: Dr Micah Nehring

Submission date: 20th of May, 2019

FORECASTING THE PRICE OF IRON ORE

A BRIEF INVESTIGATION USING REGRESSION

MINE4123

BY JAMES OMUNDSEN

SUPERVISOR: DR MICAH NEHRING

MON 20TH OF MAY 2019

STATEMENT OF ORIGINALITY

I formally declare that this investigation is my own original work and that all sources of information utilised in it have been properly acknowledged.

EXECUTIVE SUMMARY

The response of mining operations to changes in commodity prices are typically limited and delayed. Prediction may enable the reduction of delay in implementing mine planning strategy to ultimately maximise value.

The purpose of this investigation was to understand the key drivers affecting the price of Iron Ore to enable forecasting. Quarterly data was retrieved for the price of Iron Ore from September 2009 to December 2018 as well as possible relevant independent variables. EViews statistical software enabled regression modelling using Autoregressive, Leading Indicator, Autoregressive Distributed Lag, Autoregressive Moving Average and simple variated forms.

This investigation determined that the best way to model the price of Iron Ore was using a variation of the Autoregressive Moving Average model with the inclusion of a leading indicator. This was determined by R^2, Akaike's and Schwartz Information Criteria. The model used the price of Iron Ore from 1 quarter ago, the error term value from 3, and Crude Steel Production in China from 9 as well as a constant. This is demonstrated with the equation,

$$IOP_t = 134.0886 + 0.566401\,IOP_{t-1} + 0.477979\,u_{t-3} - 0.524117\,CSC_{t-9} + e_t.$$

It was assumed that this model would be most suitable for forecasting and the price of Iron Ore in March 2019 was predicted to be US$64.42/t. In reality it was US$85.75, representing a total error of negative US$21.33/t, likely caused by the tailings dam failure in Jan 2019 at Vale's Brumadinho mine in Brazil. This caused an immediate cease in production and shock loss of market supply and demonstrates the greatest weakness in the models from this investigation; they work best in stationary times and can't anticipate major shocks to the market.

This Autoregressive Moving Average model requires data from at least 1 quarter previously but this is not sufficient time to implement effective mine planning strategies. It is recommended that further investigations focus on understanding how to best model the price of Iron Ore at lags of 4 quarters onwards. It is also recommended that further investigations reduce the number of observations used to model the data and thus give models more observations to compare forecasts with reality. Consideration could also be given to market desperation with a variable expressing the net of supply and demand in the market. The relevant data, assumptions, model forms, decision rules, comprehensive results, analysis and remaining conclusions and recommendations can be found within this report.

ACKNOWLEDGMENTS

I would like to acknowledge Dr Micah Nehring whose guidance has been greatly valued to this investigation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

1. INTRODUCTION

Preconceived notions are that the purpose of mining, which is to extract minerals from or beneath the Earth's surface, is to provide the raw materials needed for the modern world. This notion is misguided, be it due to sentiment or lack of clarity. In reality, mining companies are solely focused on the idea of value with their ultimate purpose being to maximise value and return it to shareholders. It should be noted that this search for value must be conducted while adhering to relevant legal, social and environmental obligations with it being possible to argue that failure to adhere to these issues could destroy value.

A Resource is defined as a mineral occurrence and a reserve, or ore, is defined as the portions of that resource that can be feasible, or profitably, extracted. To determine what is ore and what is not, a resource is first separated into blocks, or smaller more manageable segments, to provide more comprehensive knowledge of the mineral occurrence. These blocks are then valued with consideration to the revenue that can be generated from the minerals within it, the cost of mining it and the cost of processing it to liberate the minerals. The block value formula is demonstrated as,

$$Block\ Value = Revenue - Mining\ Cost - Processing\ Cost. \qquad \textit{Equation 1}$$

The mining cost, \$, is determined by the mining cost per tonne of material denoted mc and the tonnes, t, of material being processed. This mining cost formula is demonstrated as,

$$Mining\ Cost = mc \times T. \qquad \text{Equation 2}$$

The processing cost, \$, is determined by the processing cost per tonne of material denoted pc and the tonnes, t, of material being processed. This mining cost formula is demonstrated as,

$$Processing\ Cost = pc \times T. \qquad \text{Equation 3}$$

The block revenue, \$, is determined by the price, \$ of the mineral denoted by P, the metallurgical recovery denoted by r, the grade of the mineral denoted by g, and the tonnes, t, of material being processed denoted by T. This block revenue formula is demonstrated as,

$$Revenue = P \times r \times g \times T. \qquad \text{Equation 4}$$

Of particular significance is the commodity price as it is beyond the control of any mining operation, in perfectly competitive markets. The price of any commodity is determined by the price that buyers and sellers are willing to trade for a unit of it in the marketplace. Although contracts with customers are negotiated in advance, clearly specifying the grade, total amount, time period and unit price, the benchmark for contract prices. The price of a commodity determines the cut-off grade or minimum grade of a block that can be extracted to cover the cost of mining it and the total amount of a resource considered as reserve or ore. This definition ultimately affects the mining methods used to extract it, the size and design of the mine as well as the expected life of the mine.

The value of resource is significantly more sensitive to fluctuations in commodity prices, of which mining companies have negligible influence over, as opposed to the operating practises utilised, of which mining companies have complete control over. Changes in commodity prices therefore represent both significant opportunities and risks to the value of resources, the feasible extraction of them and ultimately the maximising of value and its return to shareholders.

Predicting the movement of commodity prices is considered incredibly difficult to the point of being beyond comprehension although the simple fact that they are not randomly determined means that there is specific and justified reasoning for them. Research into long term movements would provide increased understanding of their key drivers and the opportunity for mining companies to reduce the delay in action and to implement more effective strategies to maximise value.

Admittedly, short term fluctuations in commodity prices are difficult to predict especially for commodities, such as gold and silver for example, which not only have industrial uses but are also used as financial instruments for investments and wealth protection in uncertain times. Such commodities with uses as financial instruments are thus significantly influenced by speculation and by institutional investors. Research into longer term commodity price movements of minerals used as industrial materials therefore represents more reasonable and useful opportunities.

The prices at which mining companies sell the commodities that they extract directly affects the profit margins that they make, the feasibility of them conducting business and ultimately their ability to maximise value and return it to shareholders. The foresight to understand when the good times and bad times can be expected will enable opportunities to be identified and seized as well as risks to be identified and mitigated.

2. AIM AND OBJECTIVES

The aim of this investigation is to develop a develop a deep enough understanding of the key drivers affecting the price of Iron Ore to enable forecasting.

The objectives involved with this investigation include;

1. Conducting a literature review,
2. Identifying key drivers affecting the price of Iron Ore,
3. Collecting relevant data associated with the identified key drivers,
4. Conducting correlation analysis to identify possible relevant lags, and
5. Conducting regression analysis to quantitatively understand the relationship between the identified key drivers and the price of Iron Ore.

3. SCOPE

The scope of this investigation includes, but is not limited to;

- Iron Ore and its associated issues,
- The price of Iron Ore between September 2009 to December 2018,
- Longer term Iron Ore prices, currently at intervals of quarterly or 3 month intervals,
- Key drivers to the Iron Ore prices which may include macroeconomic factors including population growth, interest rates, inflation and GDP,
- Current trends in the supply and demand of Iron Ore, and
- Regression analysis and modelling methods.

This investigation is more inclined to researching Iron Ore because of its significance to the Australian mining industry and its consistent relevance to global markets as opposed to other commodities with sporadic relevance. Iron Ore is used for industrial purposes as opposed to other commodities with uses as financial instruments and so has consistent market relevance with less influence from speculation. Foreign exchange rates are considered to be out of the scope of this investigation.

4. ASSUMPTIONS

The major assumptions relevant to this investigation are that:

- The data retrieved and utilised in this investigation are correct observations,
- Data is available instantly at the end of the quarter,
- Quarterly values for independent variables are relevant to monthly averages of the price of Iron Ore for quarters following,
- The relationships between variables determined from modelling the data set are unchanged beyond the timeframe of the data set,
- Linear relationships between independent variables and the price of Iron Ore are the most suitable way to test relationships,
- Quarterly data is frequent enough to be relevant to longer term trends but infrequent enough to be affected by short term speculation,
- All data Iron Ore data relates to 62% Iron content,
- Lags of independent variables that maximise correlation with the price of Iron Ore are best lags of variables for regression modelling, and
- Sales data for the four major Iron Ore producers are representative of the total sales in the world; Vale, BHP, Rio Tinto, and Fortescue,

5. LITERATURE REVIEW

5.1 WHAT IS IRON ORE?

Iron, denoted Fe from Latin *ferrum*, is a metallic element. It is the fourth most abundant element in the Earth's crust and the most abundant by mass, constituting approximately 5.6% of the weight of the Earth's crust (Geoscience Australia, 2016). Iron Ore refers to rocks which contain Iron and from it can be feasibly extracted. The two most prominent rocks containing Iron include Hematite, FE2O3, and Magnetite, Fe3O4.

5.2 WHAT IS IRON ORE USED FOR?

Iron ore is used to produce Iron which is predominantly used to make steel (King, 2018). The construction sector represents the largest share of steel use with an estimated share of 50% of global steel production (Department of Industry, Innovation and Science, 2018). Mechanical machinery represents an estimated share of 16% of global steel production, the automotive sector represents an estimated share of 13%, the bicycle sector represents an estimated share of 5%, the computer sector representing an estimated share of 4% and the kitchen appliance sector representing an estimated share of 2%.

5.3 HOW IS IRON ORE PRICED?

In recent years Iron Ore has entered the dynamic marketplace and is now bought and sold at the spot price. The spot price is the price at which something can be bought and sold with immediate delivery (Investopedia, 2018). It is calculated by using the prices quoted from benchmark indices and futures markets (Market Index, 2018). Benchmark indices collect and assess industry data, which can be difficult to retrieve since transactions between buyers and sellers are private with varying currencies and grades. Futures markets are places for buyers to buy or sellers to sell something at a price and date previously agreed (Investopedia, 2018).

5.4 IRON ORE PRICE SPECULATION

Iron Ore is an industrial material although its recent entrance into spot pricing in the marketplace has introduced the commodity to speculation. Before spot pricing was introduced mining companies would enter agreements directly with their buyers who were steel mills. The entrance of Iron Ore to the dynamic financial marketplace has introduced financial intermediaries into the chain. These intermediaries have the power to manipulate the price of the commodity, for their own objectives, which can have industry wide ramifications (McHugh, 2016). That being said, compared to other commodities, such as gold and silver, which are heavily used as financial instruments, Iron Ore is less susceptible to speculation and its origins are far less varied.

It must be said that speculation is less of an issue the longer the time frame; speculators that buy and sell with the intention of reversing their position later when they predict that the price will move operate in short time frames, on daily or weekly time frames. As the time frame is extended, for example to quarterly intervals, speculation become less of an issue with longer term trends dictating the movement in the price.

5.5 SIGNIFICANT PRODUCERS

Perfect competition is defined as a hypothetical market where competition is as high as it can be, maximising the benefit to consumers (Economics Online, 2018). There are several characteristics integral to perfect competition, although in reality very few industries actually achieve them all. In particular, very must be many firms in the market with easy access to entry and no firm, on its own, can influence the market price. With respect to the Iron Ore sector, this does not really apply as the industry is dominated by a select few, large mining companies. These include BHP Billiton, Vale, Rio Tinto and Fortescue Metals Group which together control more than 70% of the seaborne Iron Ore market (Investopedia, 2018).

5.6 BREAKEVEN COST OF PRODUCTION

Due to the size of the large Iron Ore mining companies they enjoy economies of scale, which is defined as a reduction in unit costs as production is increased. Reported in 2017, small scale

Iron Ore mining companies in Australia had a breakeven cost of production of $US60 per tonne (Lannin, 2017). This is in compared to the breakeven cost of production for the large four Iron Ore mining companies who reportedly, at the same time, had a production cost of $US30 per tonne. Although these numbers will likely have changed in recent times, with data difficult to obtain, the cost of production is trending down due to lower Iron Ore Prices, causing operations to increase efficiency, and innovations to the industry, such as autonomous vehicles.

5.7 CURRENT IMPORT AND EXPORT TRENDS

Australia is the largest exporter of Iron Ore in the world, exporting an estimated share of 818 million tonnes in the 2016-2017 period, representing an estimated share of 53% of global Iron Ore exports (Department of Industry, Innovation and Science, 2018). The second largest exporter, in the 2016-2017 period, was Brazil with an estimated share of 24% of global iron ore exports. The subsequent Iron Ore exporting nations each only represent an estimated share of 3-4% of global Iron Ore exports with them totalling an estimated share of 23%.

China is the largest importer of Iron Ore in the world, importing an estimated share of 67% of global Iron Ore imports in the 2016-2017 period (Department of Industry, Innovation and Science, 2018). The European Union is the second largest importer, in the 2016-2017 period, with an estimated share of 10% of global Iron Ore Imports with Japan, South Korea and the rest of the world representing an estimated share of 8%, 5% and 10% respectively.

5.8 FORECASTED TRENDS

China's significance to the Iron Ore industry means that changes to its supply and demand have extensive consequences to the industry. Its total imports in the five months up to May 2018 was 446 million tonnes, approximately the same as the previous intervals (Department of Industry, Innovation and Science, 2018). Declines to steel production have guided forecasts from the Department of Industry, Innovation and Science by the Chief Economist that China's Iron Ore imports will slow decline at 0.6% annually.

India is predicted to have consumption larger than production starting from 2019 (Department of Industry, Innovation and Science, 2018) meaning that it will become a net importer of Iron Ore. It can therefore be expected that India will become a more significant player in the global Iron Ore industry.

Global export volumes are predicted to increase by an estimated 4.3% in 2018 and 1.9% in 2019 (Department of Industry, Innovation and Science, 2018). This is due to the opening of projects in Brazil as well as trend of Australian producers showing no sign of slowing down.

5.9 POSSIBLE KEY DRIVERS

Possible key drivers for the price of Iron Ore that have been identified so far include iron ore supply, iron ore demand, crude steel production, the growth rate of real GDP, the growth rate of population, the price of oil, the price of copper and the US Federal Funds Rate.

5.9.1 SUPPLY

As the big firms reduce their unit cost of production, muscling out the firms with higher costs that cannot sustain falls in the spot price of Iron Ore, the firms that are left are apparently trying to increase their share of global production whilst demand seems to be softening. With predictions that the global production of Iron Ore will increase in the coming years, an increase in supply will cause a decrease in the price of Iron Ore.

5.9.2 DEMAND

Volatility in commodity prices seem more susceptible to shocks to demand than to supply (Coates, et al., 2011). With particular respect to the Iron Ore industry, it is likely that this argument holds true. Nations with high steel production, particularly China who produced an estimated share of 49% of global steel production in 2017, have extreme significance to the Iron Ore industry and so also the price of Iron Ore (Department of Industry, Innovation and Science, 2018).

5.9.3 CRUDE STEEL PRODUCTION

Steel is a combination of iron and carbon and is widely considered as one of the most important materials for engineering and construction (World Steel Association, Unknown). Data from the World Steel Association stated that China accounted for approximately 51.87% of the total Crude Steel Produced in the world.

It is suspected that the significance of China's Crude Steel Production with respect to the world, and so it's need for Iron Ore, mean that it may be an indicator for the price of Iron Ore.

5.9.4 REAL GDP GROWTH

Gross domestic product GDP, is significant indicator of the state of a nation's economy (Investopedia, 2019). It is defined as the sum of all goods and services produced in a given time period. Real GDP is defined similarly but is adjusted for inflation. The growth rate of Real GDP is defined as the change in Real GDP between periods.

It is suspected that there may be a relationship between the Real GDP Growth Rate of China and the price of Iron Ore.

5.9.5 OIL PRICE

The price of Iron Ore is likely very reliant on the price of Oil, in the short run (Bazhanov, 2018). This is due to the need to transport Iron Ore to and from locations overseas. The importance of the price of Oil to the price of Iron Ore will likely change in the future as the world begins to embrace renewable energy and rely less on oil as a fuel for transportation.

5.9.6 COPPER PRICE

Copper is an abundant metal which is malleable, resistant to corrosion and an efficient conductor (Geology.com, Unknown). Its favourable characteristics have enabled it to be utilised in many applications including construction, electrical and electronics, transportation equipment, consumer products and machinery. The modern world's reliance on copper has enabled its price to be an indicator for the state of an economy (Oil Price, 2009).
It is assumed that there may be some correlation between the price of copper and the price of iron ore considering that both commodities are so significant to growth.

5.9.7 FEDERAL FUNDS RATE

There is an observed negative relationship between interest rates and investment (Hambur, et al., 2018). This means that as interest rates increase, investment decreases. The intuition is that it becomes more expensive to borrow capital so firms are less willing to seek debt. Mining companies require large capital investments, usually financed through debt due to the tax shield benefit, to conduct operations, typically when beginning operations or acquiring others. With respect to the Iron Ore industry, there is likely a relationship between the interest rate on debt and the price of Iron Ore. The federal funds rate of the USA will likely be the most relevant considering the reliance of the rest of the world on it as sources of credit.

5.9.8 POPULATION

As population increases, the need for housing and accommodation increases. In countries with significant population sizes that still continue to grow, such as China, this demand for shelter causes an increase in construction particularly for cities with high population density. There is likely a relationship between the growth of population and the demand for steel. This then represents an increase in the demand for Iron Ore which will likely cause an increase in the price of Iron Ore.

5.9.9 CURRENT GLOBAL CONDITIONS

Global GDP is predicted to increase at approximately 4% in the next two years (Department of Industry, Innovation and Science, 2018). This can be due to strong labour markets and increased confidence in the USA economy. The trade tensions caused by the USA with its current trade partners represents a substantial risk to these growth predictions. Other risks to this growth in GDP include the depreciation of the Chinese currency, prolonged negotiations between Britain and the European Union with respect to the terms of its exit and subsequent trade arrangements and also the financial weaknesses in emerging economies.

5.9.10 IRON ORE PRICE MODELLING

Modelling of any commodity is incredibly difficult. Public access to those that have been created is therefore limited. Although various sources do attempt to predict commodity prices, their framework and particulars are not freely accessed. With respect to Iron Ore, this lack of public access to models is sustained.

6. REGRESSION

6.1 BACKGROUND

Regression analysis is a statistical method with the ability to estimate the relationship between certain variables. Simple Linear Regression refers to a model with a single explanatory variable. Multiple Linear Regression refers to a model with multiple explanatory variables.
Multiple Linear Regression is defined by the equation,

$$y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i. \qquad \text{Equation 5}$$

6.1.1 ASSUMPTIONS OF MULTIPLE LINEAR REGRESSION

The assumptions associated with Multiple Linear Regression include that;

- The value of the explained variable, y_i, can be explained by the explanatory variables, x_i, such that $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i, i = 1, \ldots, N,$ Equation 6
- The expected value of the random error, , is 0,
  $E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}, E(e_i) = 0,$ Equation 7
- The variance of the random error, e, is $var(e) = \sigma^2 = var(y)$,
- The covariance between any pair of random errors is $cov(e_i, e_j) = cov(y_i, y_j) = 0$, and
- The values of each x_{ik} are not random and are not exact linear functions of other explanatory variables.

6.1.2 LEAST SQUARES ESTIMATION

Least squares estimation involves finding a line which fits data the best. This is measured by squaring the distance between data points and the line and summing them all up. The line which creates the smallest total area is considered the best fit.
With Multiple Regression, we minimise the sum of squares function, $S(B_1, B_2, B_3)$, for an equation,

$E(y_i) = B_1 + x_{i2}B_2 + x_{i3}B_3$, with $S(B_1, B_2, B_3) = \sum_{i=1}^{N}(y_i - E(y_i))^2$ Equation 8

$$= \sum_{i=1}^{N} (y_i - B_1 - x_{i2}B_2 - x_{i3}B_3)^2. \qquad \text{Equation 9}$$

6.1.3 GAUSS-MARKOV THEOREM

The Gauss-Markov Theorem states that if the assumptions of Multiple Regression hold, then the least squares estimators are the best, linear, unbiased estimators of the parameters.

6.1.4 TESTING THE SIGNIFICANCE OF EXPLANATORY VARIABLES

Hypothesis testing is conducted to determine the significance of explanatory variables in modelling. To determine whether there is evidence relationship between the explained and an explanatory variable, the null hypothesis that the coefficient of an explanatory variable is 0. This is against the alternative hypothesis that the coefficient of an explanatory variable is not 0. The test statistic to determine this is the level of significance, $\propto$. If the p-value returned from modelling is smaller than the level of significance, the null hypothesis is rejected. If the p-value is larger than the level of significance, the null hypothesis is not rejected.
This investigation will use a level of significance of 0.10, $\propto = 0.10$. The framework for hypothesis testing of variables is demonstrated as,

Hypothesis:	$H_0 = 0, H_1 \neq 0$
Decision Rule:	Reject H_0 if p-value $> \propto = 0.10$ level of significance
Test Statistic:	p =
Decision:	If p-value$>\propto$, reject. If p-value$<\propto$, do not reject.

6.2 MODELS

The models used in this investigation include the Autoregressive model, Autoregressive Moving Average model, and the Autoregressive Distributed Lag model. Variations of these models may also be used.

6.2.1 AUTOREGRESSIVE, AR

The Autoregressive AR model is formed under the notion that past values of the dependent variable, in this investigation it is the price of Iron Ore, can be used to determine later values. The form of the AR model is as follows,

$$y_t = a_0 + a_1 y_{t-k} + u_t, where\ a_0 = constant, k = lag, u_t = error,$$

Equation 10

$$IOP_t = a_0 + a_1 IOP_{t-k} + u_t, where\ a_0 = constant, k = lag, u_t = error.$$

Equation 11

Modelling using the AR form will be conducted with varied lags of the price of Iron Ore.

6.2.2 MOVING AVERAGE, MA

The Moving Average MA model is formed under the notion that errors across time may be correlated. The form of the MA model is as follows,

$$u_t = b_1 e_{t-k} + e_t, where\ u_t = error, k = lag, e_t = error.$$ Equation 12

6.2.3 AUTOREGRESSIVE MOVING AVERAGE, ARMA

The Autoregressive Moving Average ARMA model IS formed under the notion that past values of the dependent variable, in this investigation it is the price of Iron Ore, as well as the errors from the AR model can be used to determine later values of the dependent variable, the price of Iron Ore. The form of the ARMA model is as follows,

$$y_t = a_0 + a_1 y_{t-k} + u_{t-k} + e_t, where\ a_0 = constant, k = lag, e_t = error,$$

Equation 13

$$IOP_t = a_0 + a_1 IOP_{t-k} + e_{t-k} + e_t, where\ a_0 = constant, k = lag, e_t = error.$$

Equation 14

Modelling using the ARMA form will be conducted with varied lags of both the price of Iron Ore and errors from the AR modelling.

6.2.4 LEADING INDICATOR

The Leading Indicator model is formed under the notion that past values of independent variables, in this investigation for example the Total Supply TSA of Iron Ore, can be used to determine later values of the dependent variable, the price of Iron Ore. The form of this model is as follows,

$$y_t = a_0 + a_1 x_{t-k} + e_t, \text{ or} \qquad \text{Equation 15}$$

$$IOP_t = a_0 + a_1 TSA_{t-1} + e_t. \qquad \text{Equation 16}$$

Modelling using the autoregressive form will be conducted with varied lags of dependent variables.

6.2.5 AUTOREGRESSIVE DISTRIBUTED LAG, ARDL

Autoregressive Distributed Lag ARDL models are formed under the notion that past values of the dependent variable, in this investigation it is the price of Iron Ore, as well as past values of independent variables, such as TSA, can be used to determine later values of the dependent variable, the price of Iron Ore. The form of the ARDL model is as follows,

$$y_t = a_0 + a_1 y_{t-k} + a_2 x_{t-l} + e_t, where\ a_0 = constant, k = lag, e_t = error$$

Equation 17

$$IOP_t = a_0 + a_1 IOP_{t-k} + TSA_{t-l} + e_t, where\ a_0 = constant, k = lag, e_t = error.$$

Equation 18

Modelling using the ARMA form will be conducted with varied lags of both the price of Iron Ore and independent variables.

6.3 CRITERIA FOR MODEL SELECTION

The models used in this investigation will be compared using three different information criteria; Coefficient of Determination R^2, Akaike's Information Criteria AIC and Bayesian Information Criterian SIC.

6.3.1 COEFFICIENT OF DETERMINATION R^2

The coefficient of determination, R^2, represents how much the independent variables in the model explain the variation in the dependent variable. It is calculated as,

$$R^2 = \frac{Sum\ of\ Squares\ Regression}{Sum\ of\ Squares\ Total} = \frac{\sum_{i=1}^{N}(\hat{y_i} - y^-)^2}{\sum_{i=1}^{N}(y_i - y^-)^2} \qquad \text{Equation 19}$$

$$= 1 - \frac{Sum\ of\ Squares\ Regression}{Sum\ of\ Squares\ Total} = 1 - \frac{\sum_{i=1}^{N} e_i^2}{\sum_{i=1}^{N}(y_i - y^-)^2}. \qquad \text{Equation 20}$$

The higher the R^2 value, the better the model explains the variation in the dependent variable. A phenomenon with this information criteria is that as more independent variables are added to the model, the R^2 value only goes up; it does not penalise for irrelevant independent variables added to the model. This is cause for concern as adding irrelevant independent variables will likely alter the coefficients and possibly significance of more relevant independent variables.

6.3.2 AKAIKE'S CRITERION, AIC

The Akaike's Criterion AIC is an information criteria used to compare models with the same dependent variable. It considers the variance of the errors which is relevant for fit, the number of observations and the number of independent variables included in the model. It is calculated as follows,

$$AIC = ln\hat{\sigma}^2 + 2\frac{p+q+1}{T}, where\ p + q = number\ of\ variables, T = number\ of\ observations. \qquad \text{Equation 21}$$

AIC is a useful criterion as it quantifies a penalty for models that do not fit the dependent variable well or use too many independent variables. Inversely, it rewards models that fit the dependent variable well and use less independent variables. The lower the AIC value, the better the model fits the dependent variable.

6.3.3 SCHWARTZ CRITERION, SIC

The Bayesian Criterion SIC is another information criteria used to compare models with the same dependent variable. It considers the variance of the errors which is relevant for fit, the number of observations and the number of independent variables included in the model. It is calculated as follows,

$$SIC = ln\hat{\sigma}^2 + \ln T \frac{p+q+1}{T}, here\ p + q = number\ of\ variables, T = number\ of\ observations. \qquad \text{Equation 22}$$

SIC is very similar to AIC and is a useful criteria as it quantifies a penalty for models that do not fit the dependent variable well or use too many independent variables. The lower the SIC value, the better the model fits the dependent variable.

The AIC and SIC models are similar although may not always when comparing models. The SIC criteria is more difficult to give low values and so is the ultimate indicator.

7. RAW DATA

7.1 SUMMARY

The dependent variable in this investigation is the price of Iron Ore. The independent variables used in this investigation include Total Sales All, Crude Steel Production China CSC, Crude Steel Production World, Real GDP Growth USA GDPU, Real GDP Growth China GDPC, Oil Price West Texas Intermediate OPW, Oil Price Brent OPB, Copper Price CP and the Federal Funds Rate FFR. A summary of the data used in this investigation can be found in Table 11. Raw Data: IOP, TSA, CSC, CSW, GDPU and GDPC, Table 12. Raw Data: OPW, OPB, CP and FFR in the appendices Raw Data.

7.2 IRON ORE PRICE, IOP

The price of Iron Ore IOP data used in this investigation has been retrieved from Market Index which is a financial portal for the Australian stock market. The units for the price of Iron Ore data is US$/t. Observations total 42 ranging from September 2008 to December 2018 at quarterly intervals. It would be more ideal for a larger data set but observations before September 2008 are yearly and thus not suitable for this investigation. As the price of Iron Ore IOP is the dependent variable in this investigation, its number of observations will generally be the upper limit for modelling unless independent variables used in modelling have fewer observations. The price history of Iron Ore is graphically demonstrated in Figure 1: Iron Ore Price History.

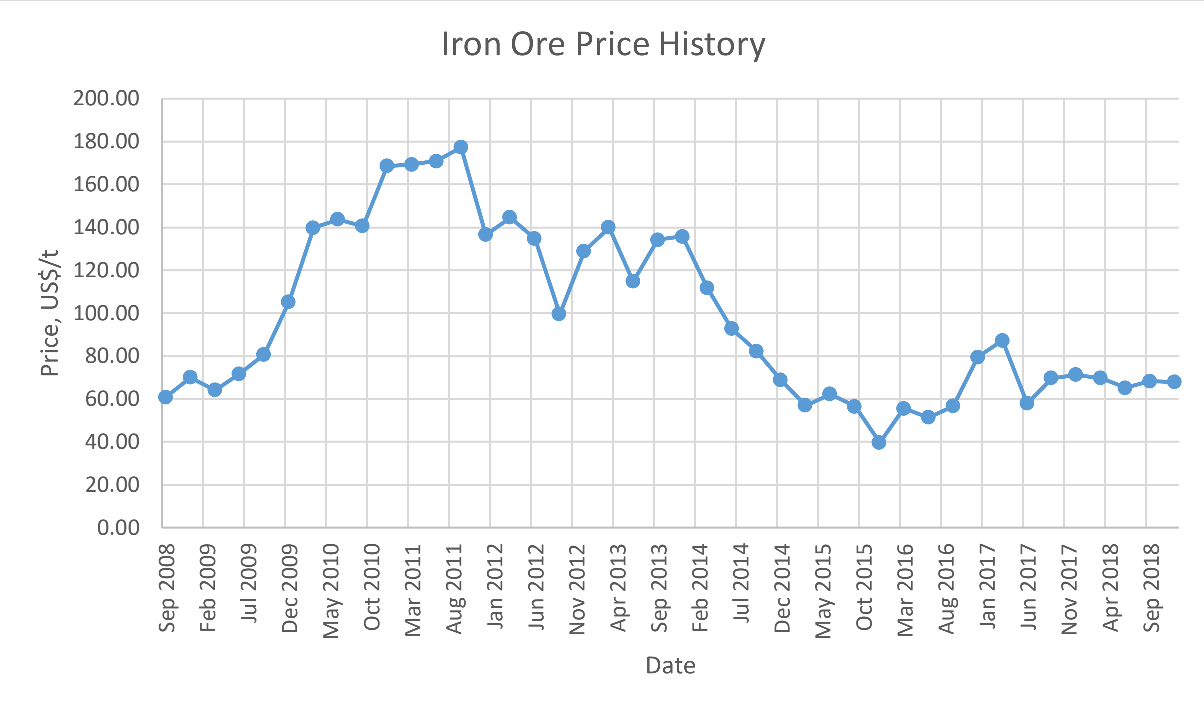

Figure 1: Iron Ore Price History

The Iron Ore price history data retrieved demonstrates a sustained increase in price from Sep 2008 to a maximum price of approximately US$180/t in September 2011. Eventually price fell back to US$60/t in March 2015 and fluctuated around approximately US$70/t until Dec 2018.

7.3 TOTAL SALES OF IRON ORE, TSA

The Total Sales All TSA data used in this investigation has been retrieved from quarterly performance reports of the four significant Iron Ore producers; Vale, BHP, Rio Tinto and Fortescue. The units of Total Sales All data is million tonne Mt. Observations total 37 ranging from December 2009 to December 2018 at quarterly intervals. It would be more ideal for a larger data set but production data from before this time is not publicly available and as this variable is the sum of the four significant producers the data set must end there. Total Sales All TSA will be an independent variable in this investigation and as the number of observations is smaller than that of the price of Iron Ore IOP, its number of observations will be the upper limit for modelling. The Total Sales All history graphically demonstrated in Figure 2: Total Sales All History.

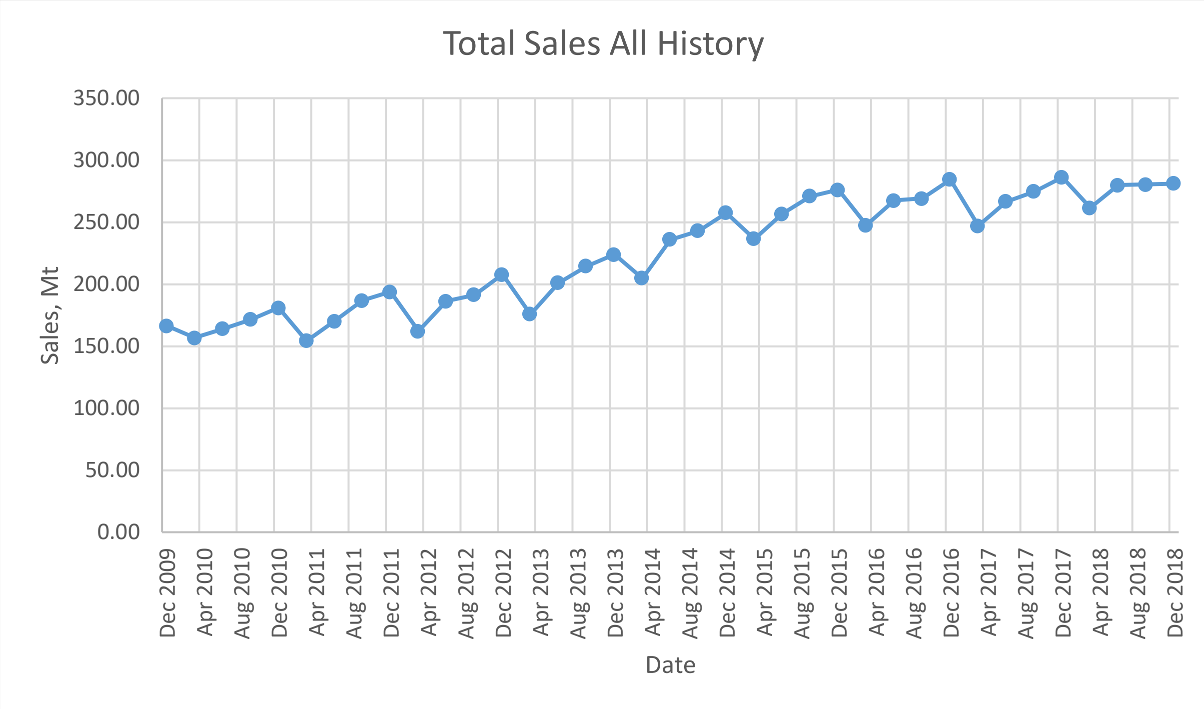

Figure 2: Total Sales All History

The Total Sales All TSA history data retrieved demonstrates a gradual increase from approximately 165Mt to approximately 250Mt with consistent drops annually in observations from March.

7.4 CRUDE STEEL PRODUCTION CHINA, CSC

The Crude Steel Production China data used in this investigation has been retrieved from World Steel Association. Observations total 43 ranging from June 2008 to December 2018 at quarterly intervals. The units of Crude Steel Production in China data is million tonne Mt. It would be more ideal for a larger data set but the significance of China to the total crude steel production in the world reduces before June 2008. As this investigation identified China as a possible significant variable in the price of Iron Ore, including observations before this time may distort the impact crude steel production in China may have using modelling. Crude Steel Production China CSC will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore

IOP observations will be the upper limit for modelling. The Crude Steel Production China history is graphically demonstrated in Figure 3. Crude Steel Production China History.

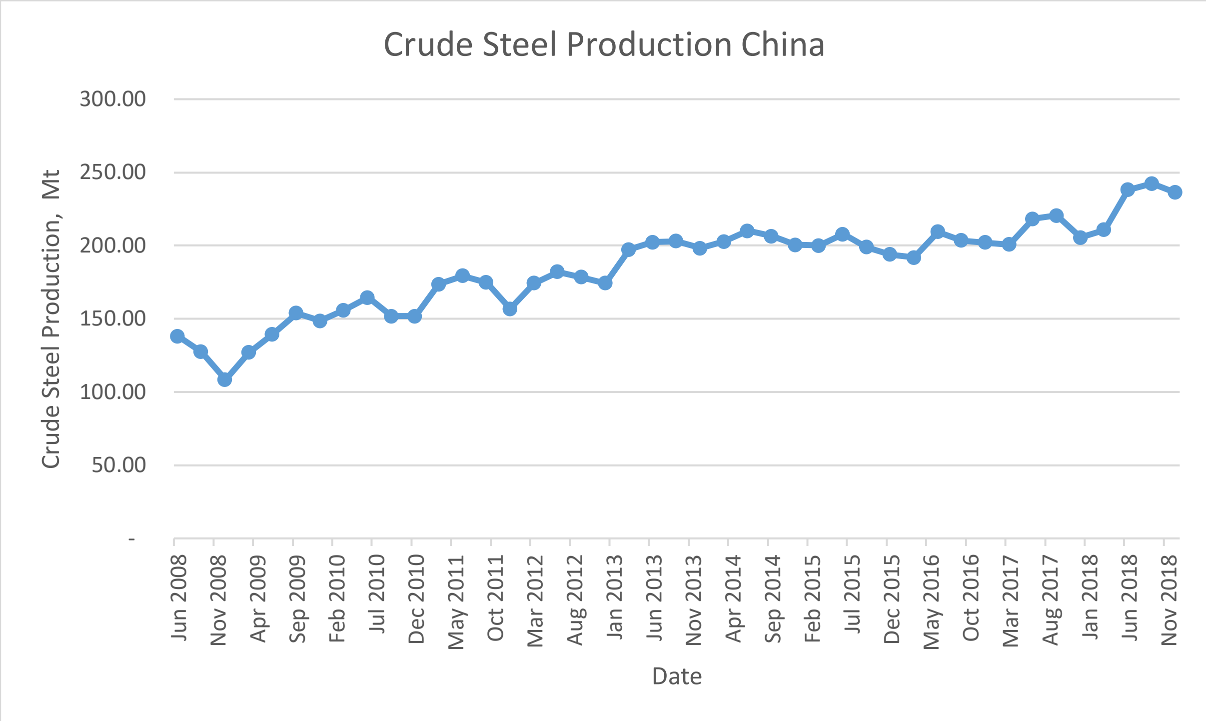

Figure 3. Crude Steel Production China History

The Crude Steel Production China CSC data demonstrates a gradual increase from approximately 138 Mt in June 2008 to 236 Mt December 2018 with general fluctuations in between.

7.5 CRUDE STEEL WORLD, CSW

The Crude Steel Production World CSW data used in this investigation has been retrieved from World Steel Association. Observations total 43 ranging from June 2008 to December 2018 at quarterly intervals. The units of Crude Steel World data is million tonne Mt. It would be more ideal for a larger data set but the significance of China to the total crude steel production in the world reduces before June 2008. As this investigation identified China as a possible significant variable in the price of Iron Ore, including observations before this time may distort the impact total crude steel production in the world may have using modelling. Crude Steel Production World CSW will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore IOP observations will be the upper limit for modelling. The Crude Steel Production China history is graphically demonstrated in Figure 4. Crude Steel Production World History.

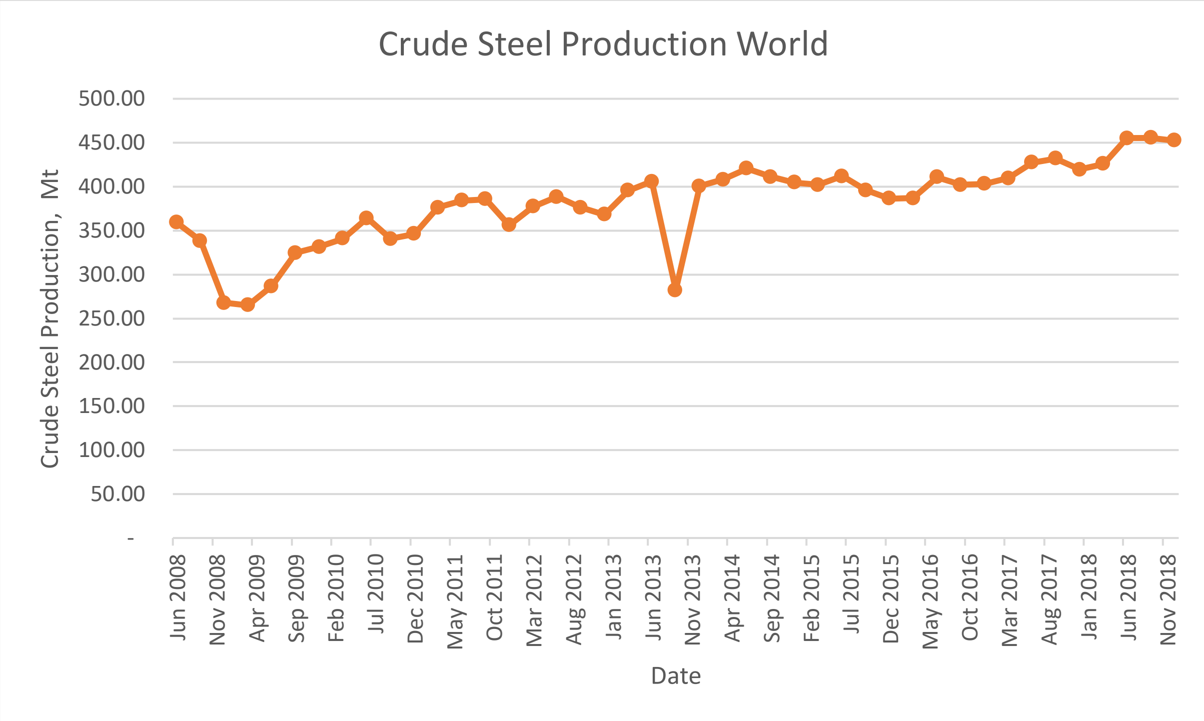

Figure 4. Crude Steel Production World History

After a steep decline from approximately 360 Mt in June 2008 to 260 Mt in November 2008, Crude Steel Production World CSW increases rather rapidly to 360 Mt July 2010 and then more gradually thereafter to December 2018. A notable decline in production is observed in September 2013 with trend production being returned to immediately the following quarter.

7.6 REAL GDP GROWTH USA, GDPU

The growth of Real GDP USA GDPU data used in this investigation has been retrieved from the OECD. Observations total 54 ranging from September 2005 to December 2018 at quarterly intervals. The units of real GDP Growth USA is percentage change %. The size of the data set is satisfactory. The growth of Real GDP USA GDPU will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore IOP observations will be the upper limit for modelling. The growth of Real GDP USA GDPU history is graphically demonstrated in Figure 5. Real GDP Growth USA History.

Figure 5. Real GDP Growth USA History

The growth of Real GDP fluctuates considerably with numerous negative growth observations. It generally fluctuates from 0.5% throughout the whole history of the data set with a distinct decrease in growth to approximately -2% in March 2009.

7.7 REAL GDP GROWTH CHINA, GDPC

The growth of Real GDP USA GDPU data used in this investigation has been retrieved from the OECD. Observations total 32 ranging from March 2011 to December 2018 at quarterly intervals. The units of Real GDP Growth China is percentage change, %. It would be more ideal for a larger data set but observations from the OECD do not continue before this date. The growth of Real GDP China GDPC will be an independent variable in this investigation and as the number of observations is smaller than that of the price of Iron Ore IOP the number of price of Iron Ore IOP observations will be the upper limit for modelling. The growth of Real GDP China GDPC history is graphically demonstrated in Figure 6. Real GDP Growth China History.

Figure 6. Real GDP Growth China History

After a rapid decrease from 2.6% in March 2011 to 1.5% in December 2011, the growth of Real GDP China generally floats around 1.75% from September 2012 to December 2018.

7.8 OIL PRICE WEST TEXAS INTERMEDIATE, OPW

The Oil Price West Texas Intermediate OPW data used in this investigation has been retrieved from Market Index which is a financial portal for the Australian stock market. Observations total 62 ranging from September 2003 to December 2018 at quarterly intervals. The units of Oil Price WTI data is US$/barrel. The size of the data set is satisfactory. The Oil Price West Texas Intermediate OPW will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore IOP observations will be the upper limit for modelling. The Oil Price West Texas Intermediate OPW history is graphically demonstrated in Figure 7. Oil Price WTI History.

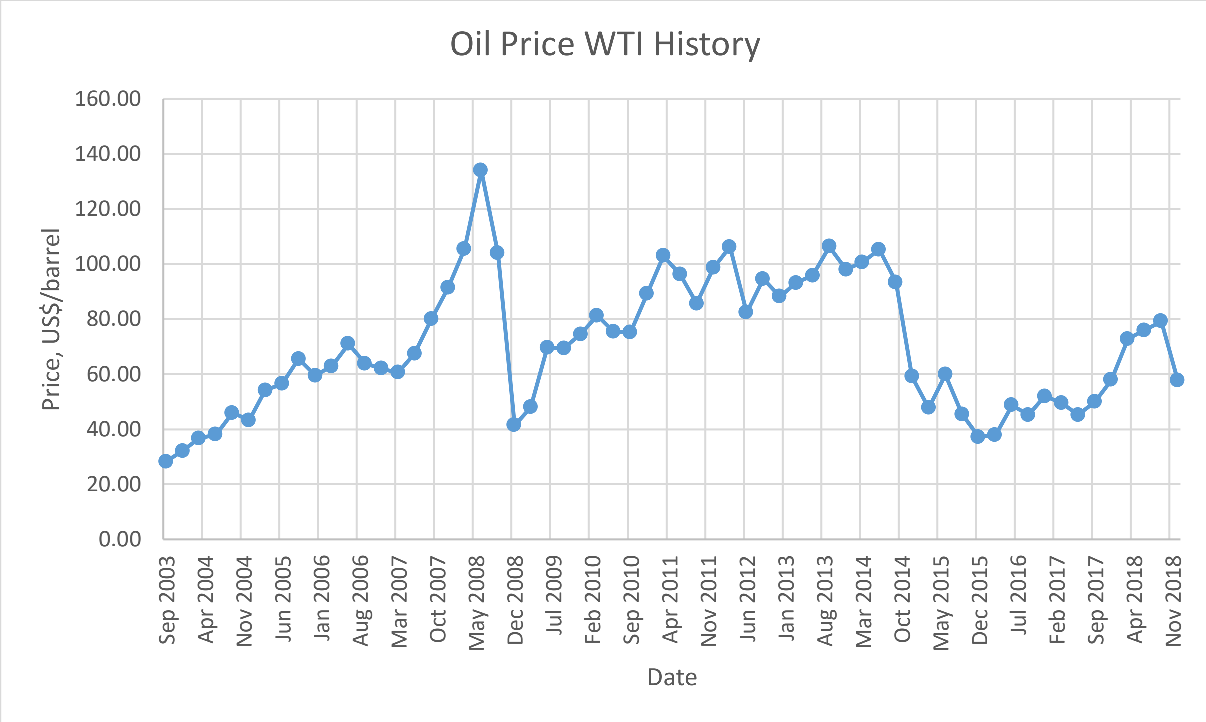

Figure 7. Oil Price WTI History

The Oil Price West Texas Intermediate OPW has an upward trend from approximately US$30/barrel in September 2003 to approximately US$105/barrel in June 2014, with a sharp decline from approximately US$140/barrel in June 2008 to approximately US$40/barrel in December 2008. Price drops significantly after June 2014, bottoming out at US$40/barrel in December 2015 then rising to US$80/barrel in June 2018 and falling to approximately US$60/barrel in December 2018.

7.9 OIL PRICE BRENT CRUDE, OPB

The Oil Price Brent Crude OPB data used in this investigation has been retrieved from Market Index which is a financial portal for the Australian stock market. Observations total 62 ranging from September 2003 to December 2018 at quarterly intervals. The units of Oil Price Brent data is US$/barrel. The size of the data set is satisfactory. The Oil Price Brent Crude OPB will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore IOP observations will be the upper limit for modelling. The Oil Price Brent OPB history is graphically demonstrated in Figure 8. Oil Price Brent History.

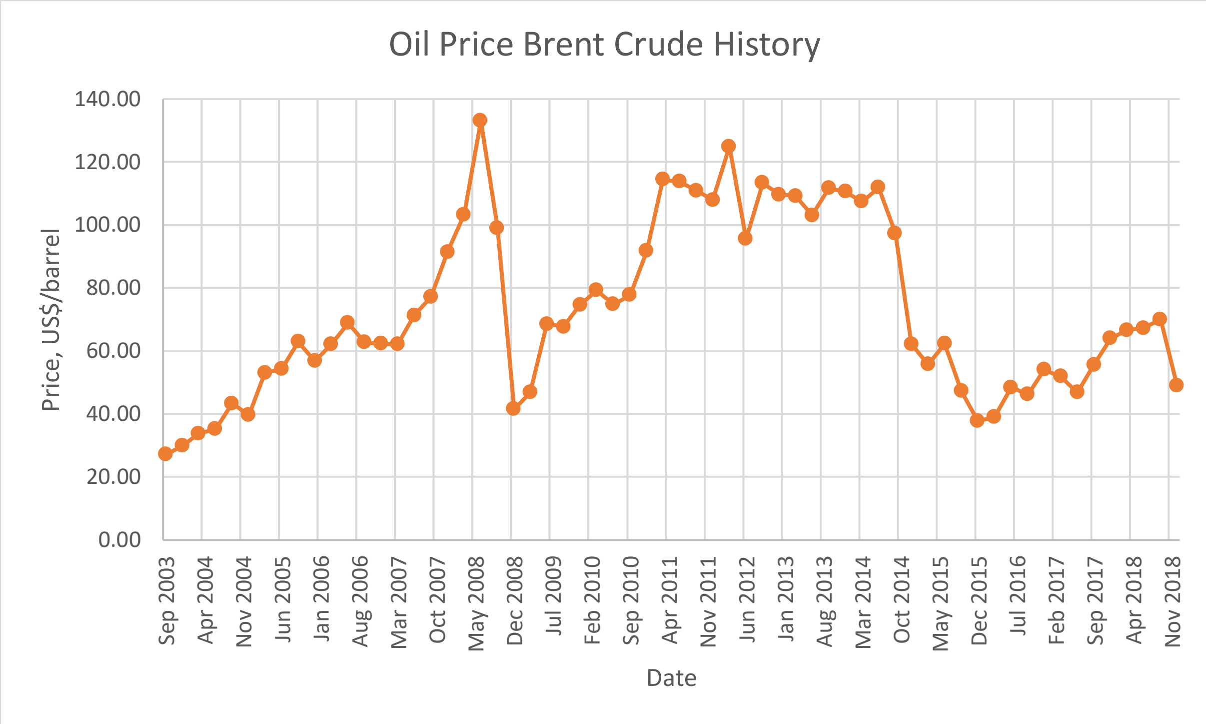

Figure 8. Oil Price Brent History

The Oil Price Brent Crude OPB has an upward trend from approximately US$30/barrel in September 2003 to approximately US$110/barrel in June 2014, with a sharp decline from approximately US$140/barrel in June 2008 to approximately US$40/barrel in December 2008. Price drops significantly after June 2014, bottoming out at US$40/barrel in December 2015 then rising to US$70/barrel in June 2018 and falling to approximately US$50/barrel in December 2018.

7.10 COPPER PRICE, CP

The Copper Price CP data used in this investigation has been retrieved from Market Index which is a financial portal for the Australian stock market. The units of Copper Price are US$/t. Observations total 62 ranging from September 2003 to December 2018 at quarterly intervals. The size of the data set is satisfactory. The Copper Price CP will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore IOP observations will be the upper limit for modelling. The Copper Price CP history is graphically demonstrated in Figure 9. Copper Price History.

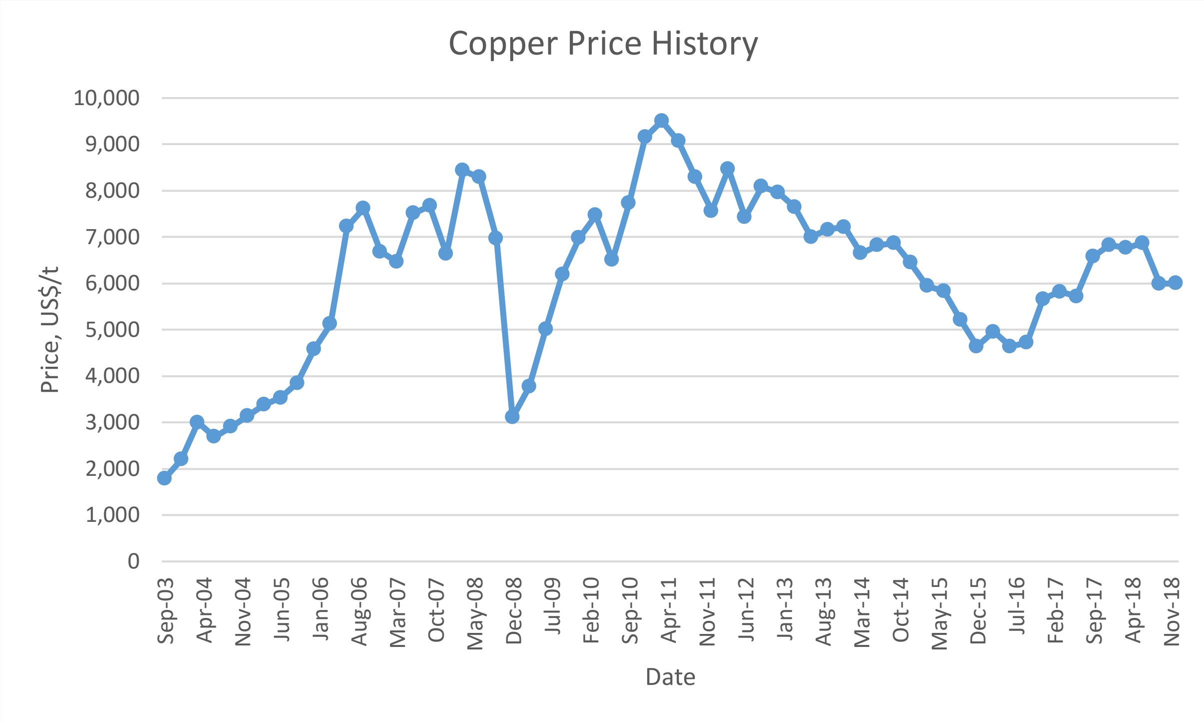

Figure 9. Copper Price History

The Copper Price CP demonstrates a steep increase in price from approximately US$2,000/t in September 2003 to approximately US$850/t in March 2008. After this it falls dramatically to approximately US$3,000/t in December 2008 but revives and increases further to approximately US$950/t in March 2011. From here it gradually decreases, bottoming out at approximately US$4,750/t in June 2016 and increases to approximately US$7,000/t in June 2018 then falls to approximately US$6,000/t in December 2018.

7.11 FEDERAL FUNDS RATE, FFR

The Federal Funds Rate FFR data used in this investigation has been retrieved from the Federal Reserve Bank of St Louis. The unit of Federal Funds Rate data is percentage %. Observations total 62 ranging from September 2003 to December 2018 at quarterly intervals. The size of the data set is satisfactory. The Federal Funds Rate FFR will be an independent variable in this investigation and as the number of observations is larger than that of the price of Iron Ore IOP, the number of price of Iron Ore IOP observations will be the upper limit for modelling. The Federal Funds Rate FR history is graphically demonstrated in Figure 10. Fed Funds Rate History.

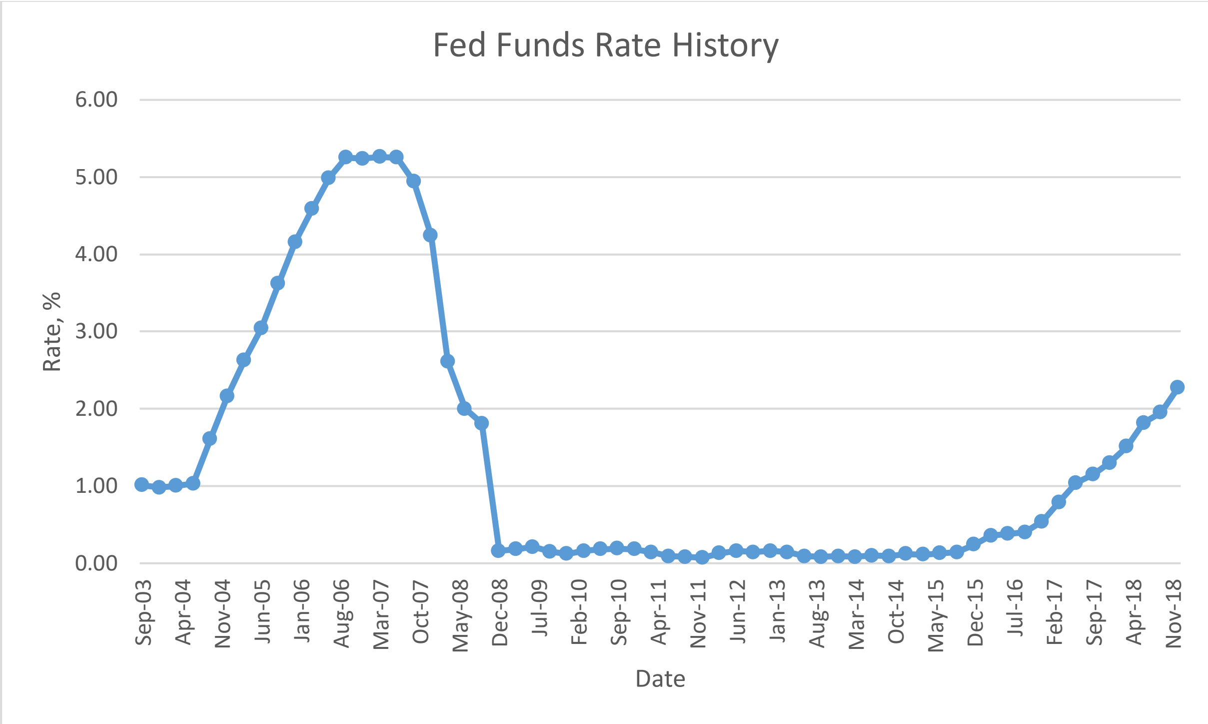

Figure 10. Fed Funds Rate History

The Federal Funds Rate FFR demonstrates a steep increase from approximately 1% in June 2004 to approximately 5.1% in September 2016. It then falls aggressively in June 2007 to 0.16% in December 2008 where it is sustained until around June 2015 where it begins to increase gradually then reaching 2.27% in December 2018.

8. METHOD

The method used to obtain data and model is as follows:

1. Retrieve, and sanitise where necessary, relevant dependent and independent variable data at quarterly periods.

2. Test correlation between the price of Iron Ore and independent variables at a lag of 1 quarter, repeating until reaching 20 quarters for each independent variable to find lag where correlation is maximised.

3. Using EViews, statistical software, import data and construct models, using the relevant Least Squares method of estimation, using all lags of independent variables where models include only 1 explanatory variable and using lags maximising correlation with the price of Iron Ore for models using more than 1 explanatory variable. Model both using a constant and without.

4. Order based on highest R^2, lowest Akaike's Informoation Criterion AIC and lowest Schwartz Information Criterion SIC. If information criterions offer conflicting orders, use Schwartz Information Criterion SIC.

9. RESULTS

9.1 CORRELATION WITH IRON ORE PRICE

Preliminary analysis was first conducted, investigating the correlation between the price of Iron Ore and the lags of independent variables. Understanding when correlation is maximised with lagged independent variables will be useful when conducting modelling. This investigation assumes that lags of variables that maximise the correlation with the price of Iron Ore will be more suitable independent variables when constructing models.

9.1.1 IRON ORE PRICE

The price of Iron ore was first correlated to its own lags. The maximum correlation observed was 0.909 at a lag of 1 quarter. After this, correlation reduced as lags increased. Due to the size of the price of Iron Ore data, the data set being tested for correlation reduced by one as the lag of the price of Iron Ore data increased by one. The value of correlation as lag against lag is visually demonstrated in Figure 11. Correlation of Iron Ore Price against Iron Ore Price.

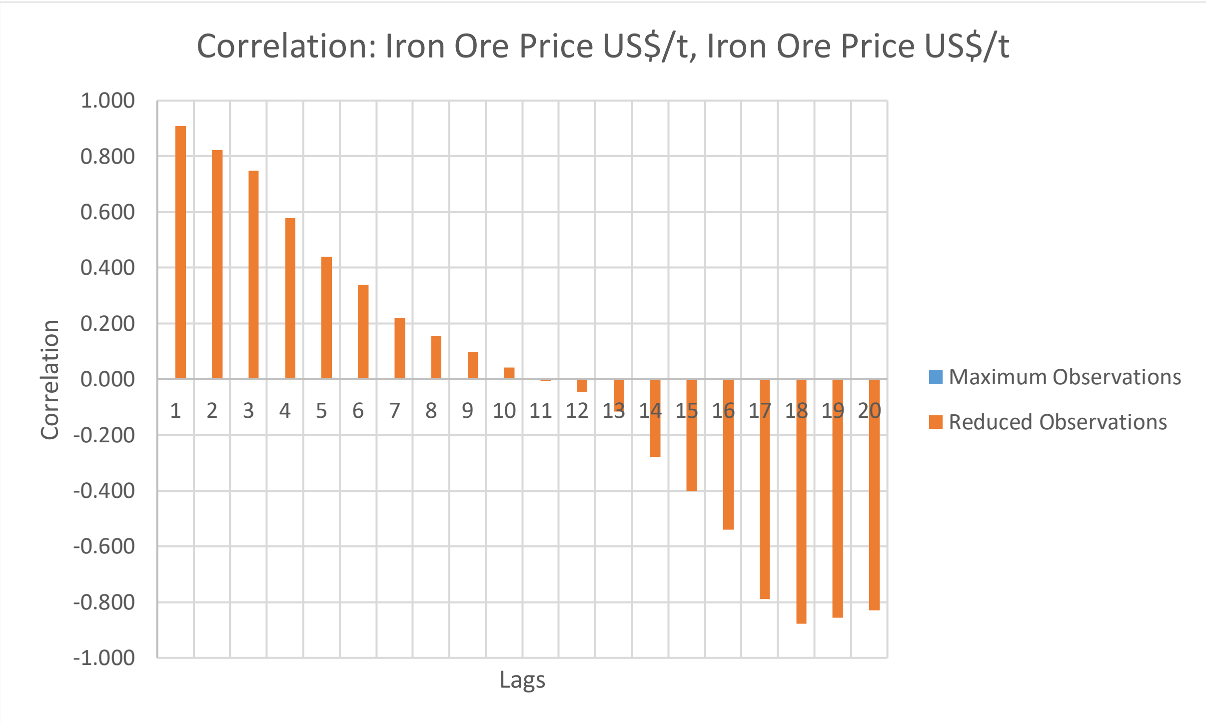

Figure 11. Correlation of Iron Ore Price against Iron Ore Price

9.1.2 TOTAL SALES ALL

The price of Iron ore was then correlated against Total Sales All. The maximum correlation observed was -0.889 at a lag of 2 quarters. After this, correlation reduced in magnitude and became positive from lags 18 to 20. Due to the Total Sales All data set being smaller than the price of Iron Ore data set, correlation was conducted first using 36 observations, 1 observation less than the full Total Sales All set, which reduced by one as the lag increased by one. The value of correlation between Iron Ore Price and Total Sales All with corresponding lag is visually demonstrated in Figure 12. Correlation Between Iron Ore Price and Total Sales All.

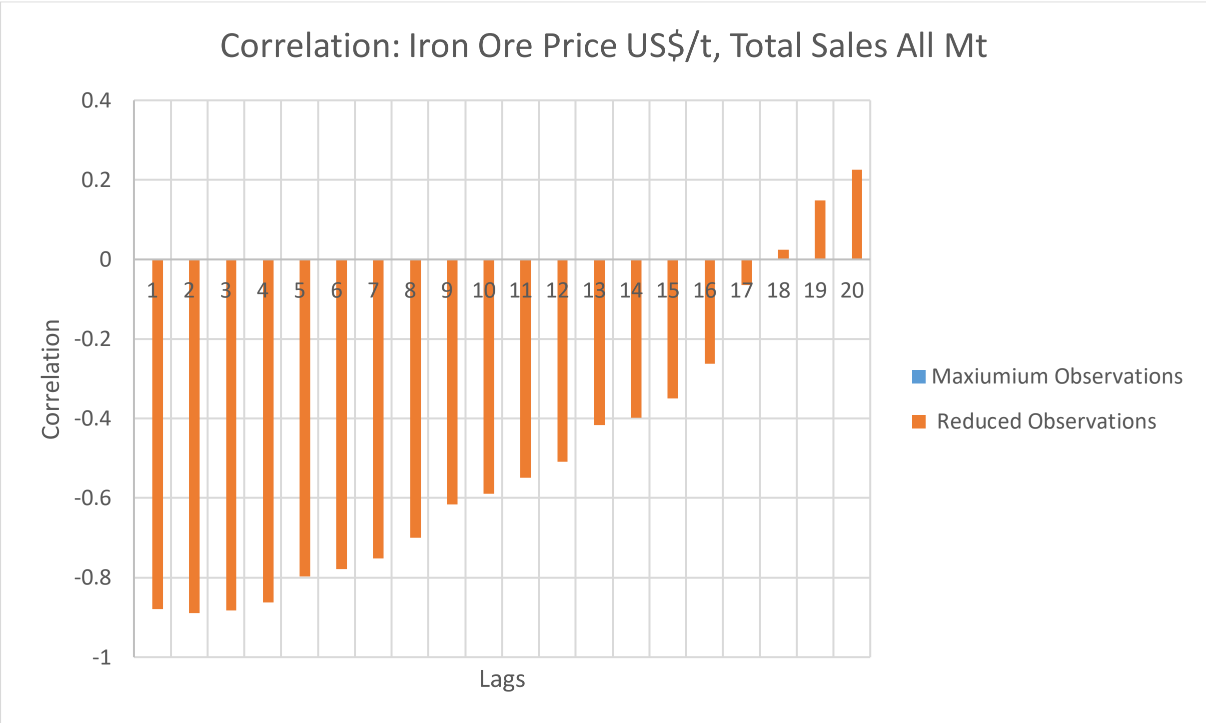

Figure 12. Correlation Between Iron Ore Price and Total Sales All

9.1.3 CRUDE STEEL PRODUCTION CHINA

The price of Iron ore was then correlated against Crude Steel Production China. The maximum correlation observed was -0.903 at a lag of 9 quarters. Correlation increased in magnitude until lag 9 where it gradually reduced until lag 15 when it became generally flat. Correlation conducted at 1 lag of Crude Steel Production China was conducted using the maximum observations of the Iron Ore Price data set of 42. With each correlation thereafter, testing size was reduced by one as the lag increased by one. The value of correlation between Iron Ore Price and Total Sales All with corresponding lag is visually demonstrated in Figure 13. Correlation Between Iron Ore Price and Crude Steel Production China.

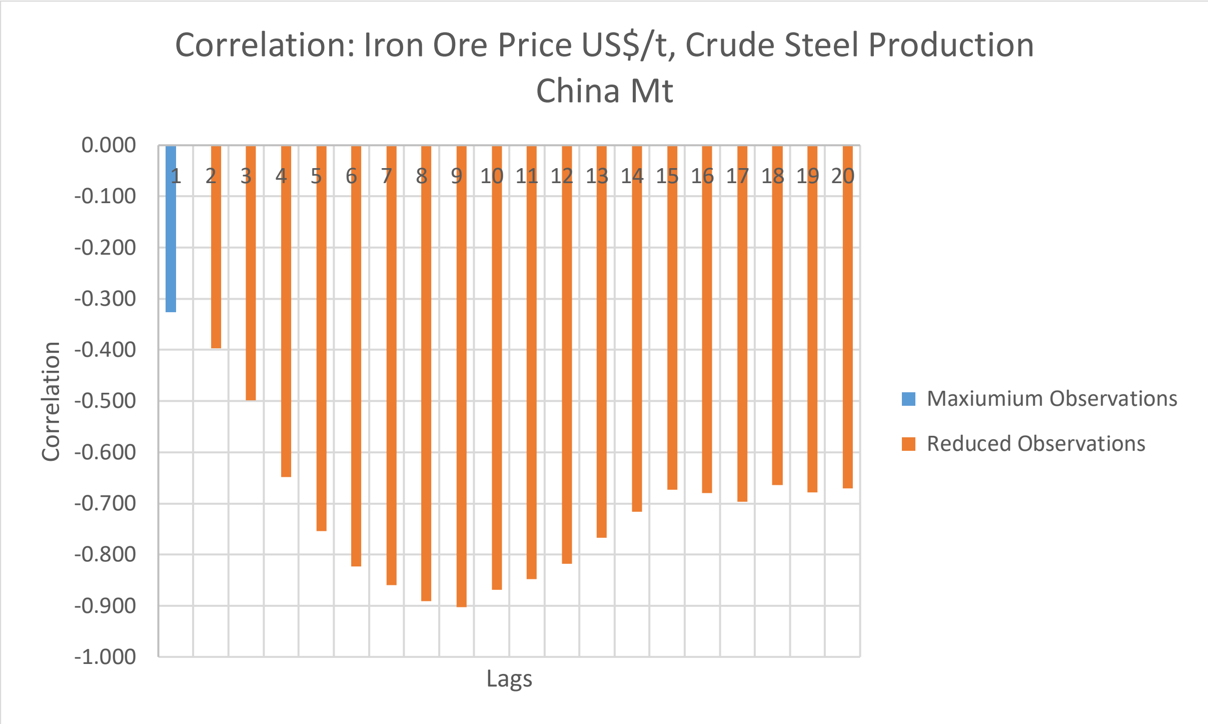

Figure 13. Correlation Between Iron Ore Price and Crude Steel Production China

9.1.4 CRUDE STEEL PRODUCTION WORLD

The price of Iron ore was then correlated against Crude Steel Production World. The maximum correlation observed was -0.719 at a lag of 8 quarters. Correlation increased in magnitude until lag 8 where it gradually reduced until lag 12 and then again increasing in magnitude at lag 17 and again reducing until lag 20. Correlation conducted at 1 lag of Crude Steel Production China was conducted using the maximum observations of the Iron Ore Price data set of 42. With each correlation thereafter, testing size was reduced by one as the lag increased by one. The value of correlation between Iron Ore Price and Crude Steel Production World with corresponding lag is visually demonstrated in Figure 14. Correlation Between Iron Ore Price and Crude Steel Production World.

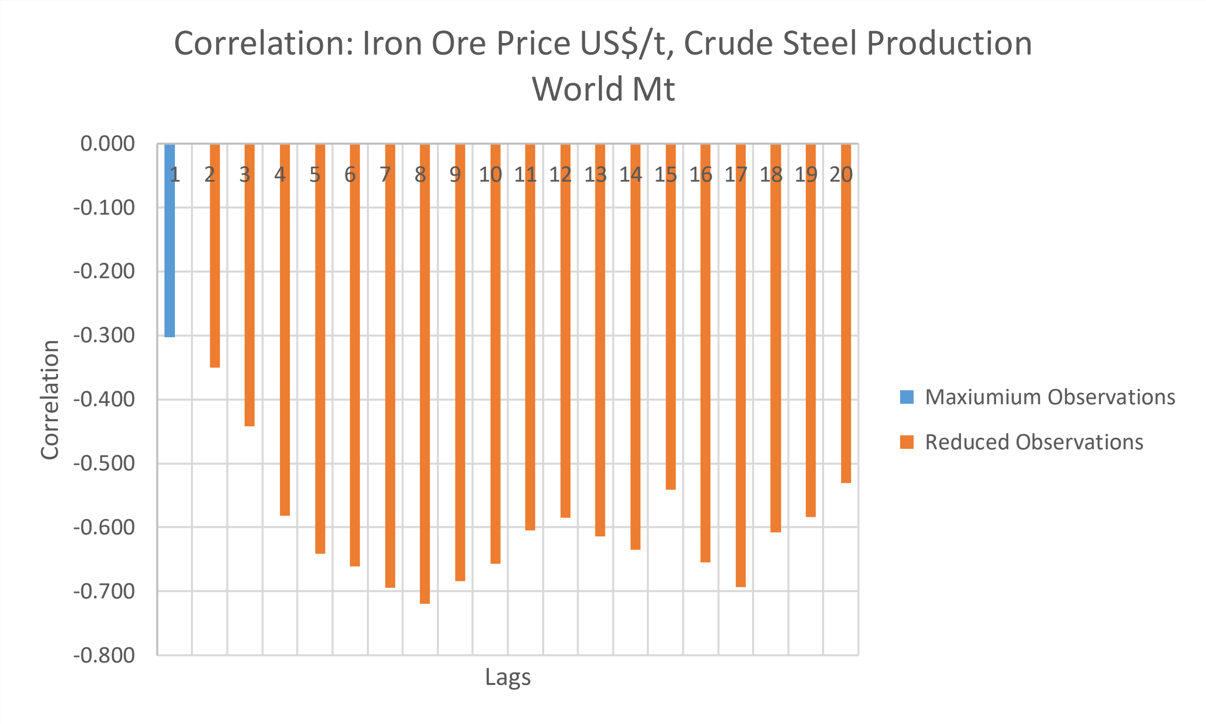

Figure 14. Correlation Between Iron Ore Price and Crude Steel Production World

9.1.5 REAL GDP GROWTH USA

The price of Iron ore was then correlated against Real GDP Growth USA. The maximum correlation observed was -0.473 at a lag of 11 quarters. Correlation was reducing but positive from lags 1 to 3 before becoming negative and increasing in magnitude until lag 11. After this it generally reduced in magnitude until lag 20. Correlation conducted until lag 12 of Real GDP Growth USA was conducted using the maximum observations of the Iron Ore Price data set of 42. With each correlation thereafter, testing size was reduced by one as the lag increased by one. The value of correlation between Iron Ore Price and Real GDP Growth USA with corresponding lag is visually demonstrated in Figure 15. Correlation Between Iron Ore Price and Real GDP Growth USA.

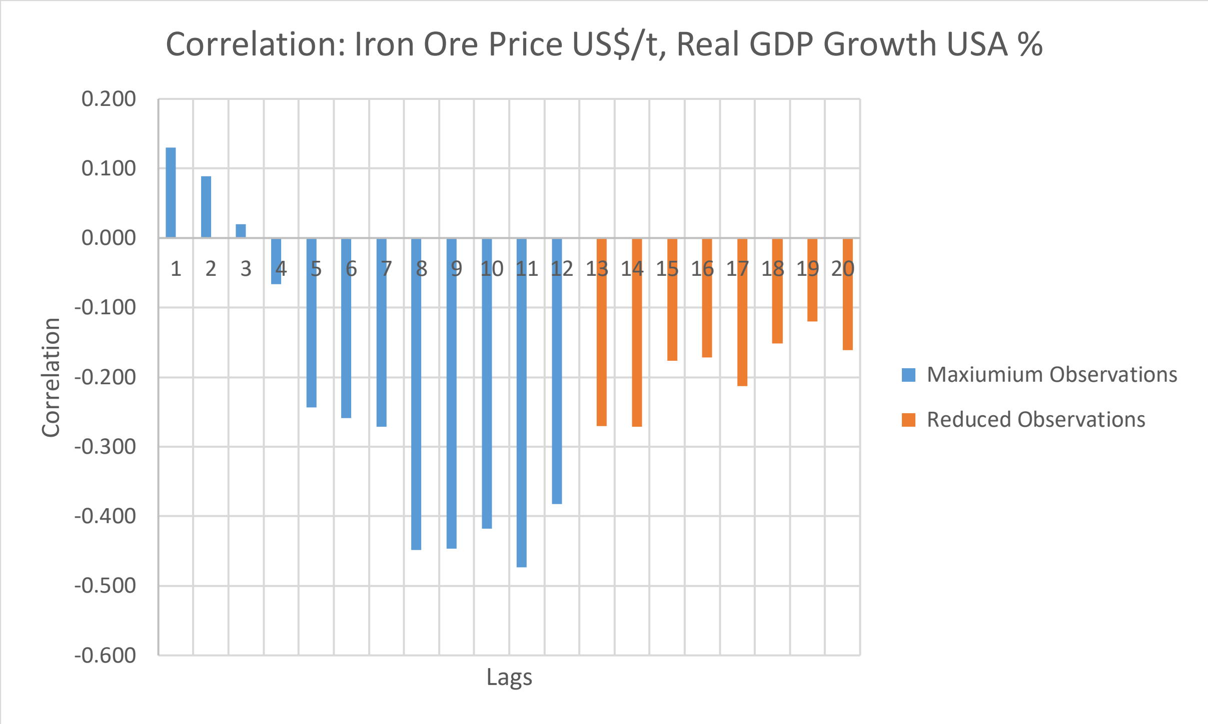

Figure 15. Correlation Between Iron Ore Price and Real GDP Growth USA

9.1.6 REAL GDP GROWTH CHINA

The price of Iron ore was then correlated against Real GDP Growth China. The maximum correlation observed was -0.654 at a lag of 1 quarter. Correlation generally reduces until lag 20 with fluctuation present. Correlation conducted at 1 lag of Crude Steel Production China was conducted using one observation less than the maximum observations of the Crude Steel Production China data set of 31. With each correlation thereafter, testing size was reduced by one as the lag increased by one. The value of correlation between Iron Ore Price and Real GDP Growth China with corresponding lag is visually demonstrated in Figure 16. Correlation Between Iron Ore Price and Real GDP Growth China.

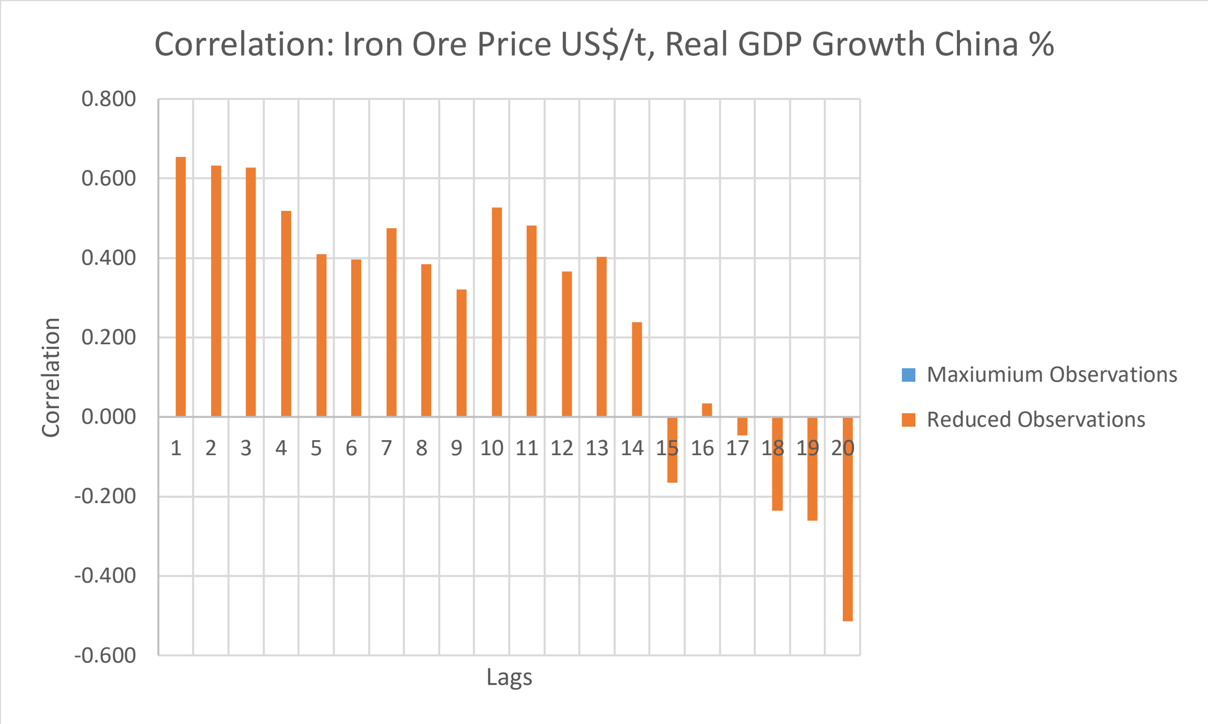

Figure 16. Correlation Between Iron Ore Price and Real GDP Growth China

9.1.7 OIL PRICE WEST TEXAS INTERMEDIATE

The price of Iron ore was then correlated against Oil Price West Texas Intermediate. The maximum correlation observed was 0.545 at a lag of 1 quarter. Correlation reduces from lag 1 until lag 5 where it is fairly consistent before becoming negative until lag 17 and then becoming more positive until lag 20. Correlation was conducted at the maximum number of observations in the Iron Ore Price data set of 42. The value of correlation between Iron Ore Price and Oil Price West Texas Intermediate with corresponding lag is visually demonstrated in Figure 17. Correlation Between Iron Ore Price and Oil Price West Texas Intermediate.

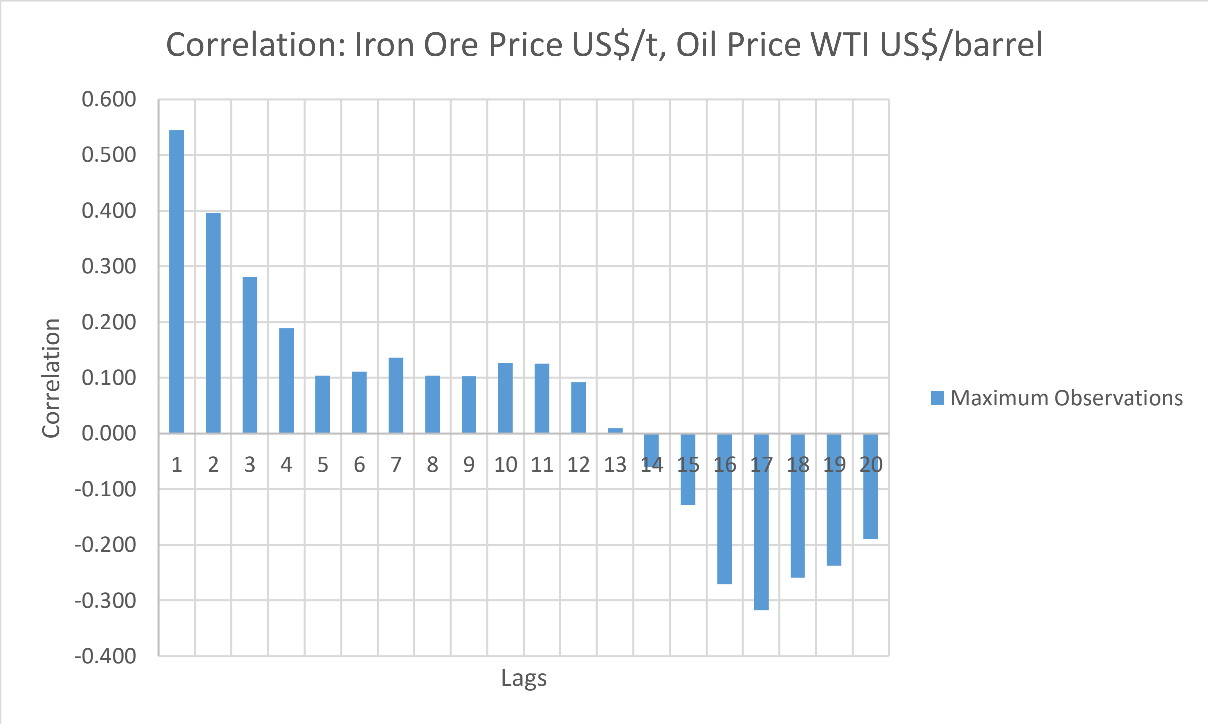

Figure 17. Correlation Between Iron Ore Price and Oil Price West Texas Intermediate

9.1.8 LAGGED OIL PRICE BRENT CRUIDE

The price of Iron ore was then correlated against Oil Price Brent Crude. The maximum correlation observed was 0.615 at a lag of 1 quarter. Correlation reduces from lag 1 until lag 10, thereafter becoming negative and reducing more but increasing in magnitude until lag 17 and then becoming more positive until lag 20. Correlation was conducted at the maximum number of observations in the Iron Ore Price data set of 42. The value of correlation between Iron Ore Price and Oil Price West Texas Intermediate with corresponding lag is visually demonstrated in Figure 18. Correlation Between Iron Ore Price and Oil Price Brent Crude.

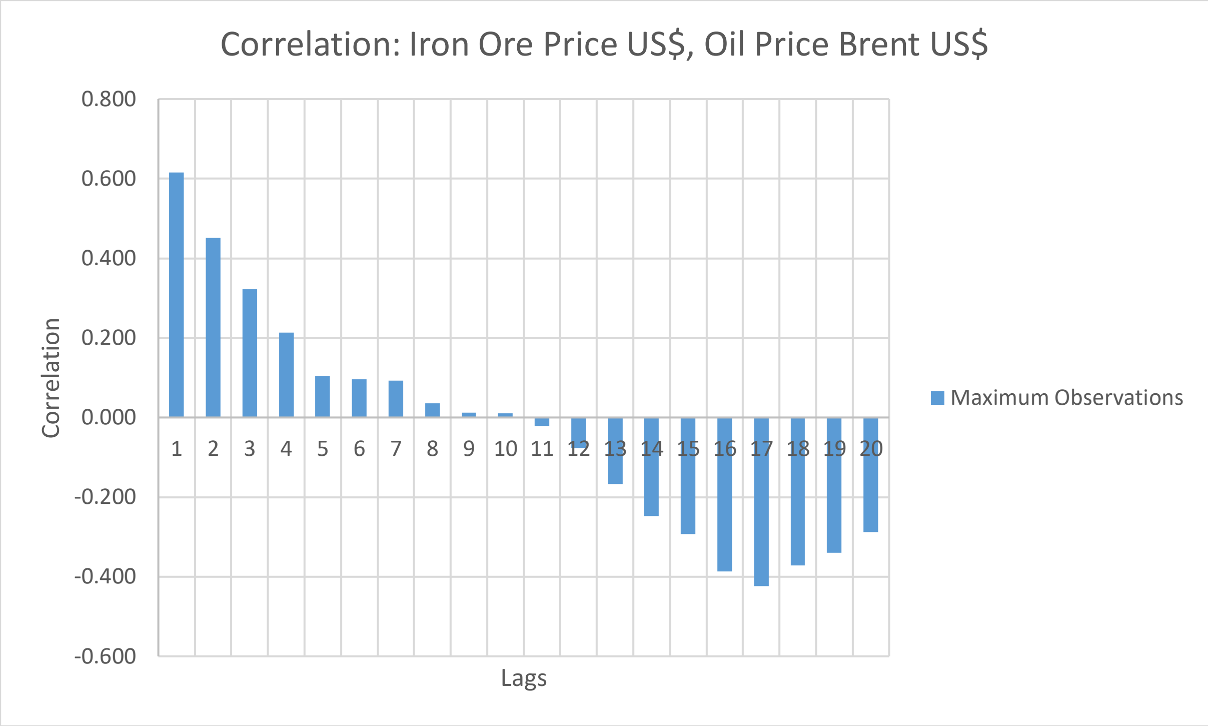

Figure 18. Correlation Between Iron Ore Price and Oil Price Brent Crude

9.1.9 LAGGED COPPER PRICE

The price of Iron ore was then correlated against Copper Price. The maximum correlation observed was 0.726 at a lag of 1 quarter. Correlation reduces from lag 1 until lag 8 where it is fairly consistent just above 0 until lag 16 where it becomes 16 and increases in magnitude until lag 20. Correlation was conducted at the maximum number of observations in the Iron Ore Price data set of 42. The value of correlation between Iron Ore Price and Copper Price with corresponding lag is visually demonstrated in Figure 19. Correlation Between Iron Ore Price and Copper Price.

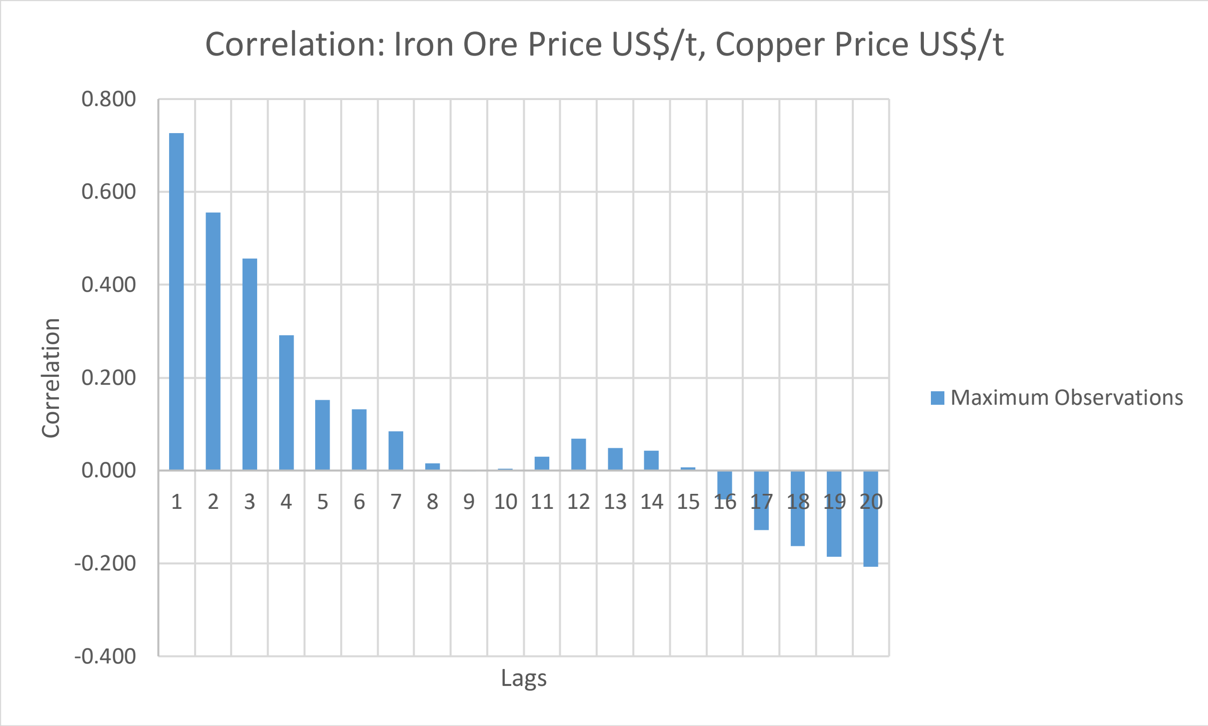

Figure 19. Correlation Between Iron Ore Price and Copper Price

9.1.10 LAGGED FED FUNDS RATE

The price of Iron ore was then correlated against Federal Funds Rate. The maximum correlation observed was 0.826 at a lag of 20 quarters. Correlation is initially negative and becomes more positive until lag 9, thereafter becoming positive and increasing in correlation until lag 20. Correlation was conducted at the maximum number of observations in the Iron Ore Price data set of 42. The value of correlation between Iron Ore Price and Federal Funds Rate with corresponding lag is visually demonstrated in Figure 20. Correlation Between Iron Ore Price and Federal Funds Rate.

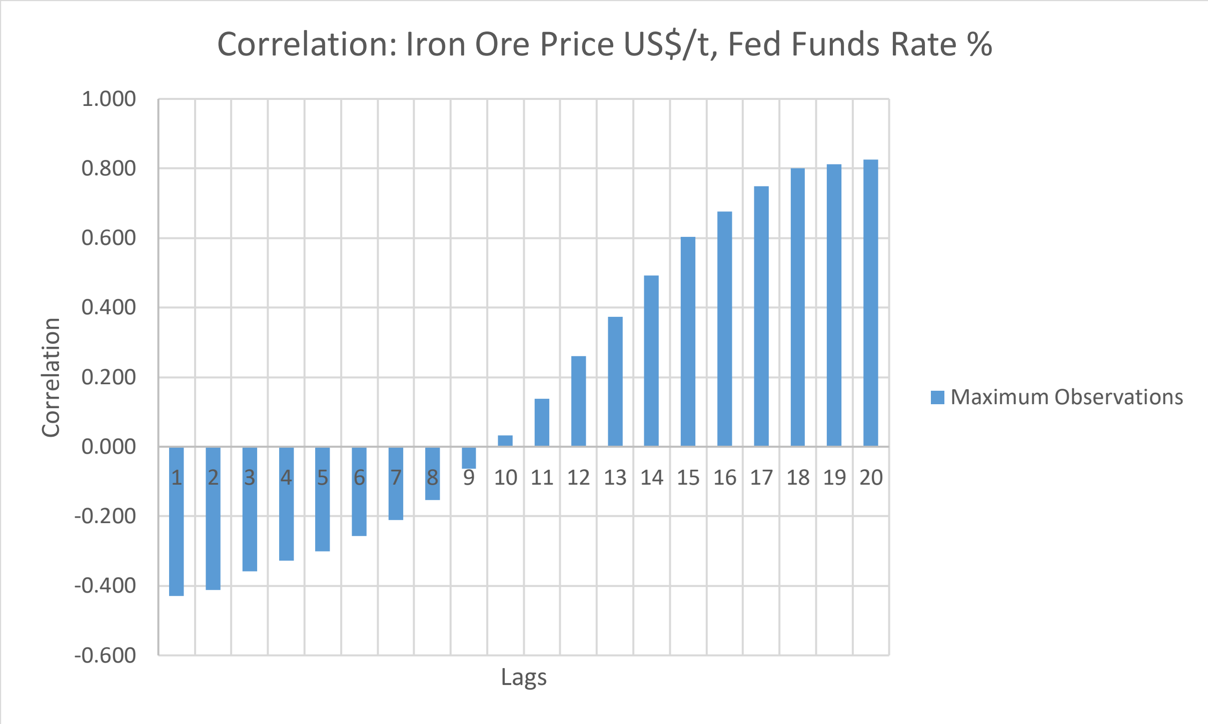

Figure 20. Correlation Between Iron Ore Price and Federal Funds Rate

9.1.11 SUMMARY

A table summarising the lag number and correlation value for each independent variable can be found in Table 1. Lag For Each Independent Variable Maximising Correlation With Iron Ore Price.

Table 1. Lag For Each Independent Variable Maximising Correlation With Iron Ore Price

Independent Variable	Lag	Correlation
IOP	1	0.909
TSA	2	-0.889
CSC	9	-0.903
CSW	8	-0.719
GDPU	11	-0.473
GDPC	1	0.654
OPW	1	0.545
OPB	1	0.615
CP	1	0.726
FFR	20	0.826

9.2 MODELLING

9.2.1 AUTOREGRESSIVE AR

Modelling was conducted using the Autoregressive AR form, both with a constant and without. The information criterions used to help indicate models that fit the data best were R^2,

Akaike's Information Criterion AIC and Schwartz Information Criterion SIC. Figure 21. R-Squared Against Lag of AR, Figure 22. AIC Against Lag Of ARand Figure 23. SIC Against Lag Of AR demonstrate how the values of these information criterion's change as the lag of the price of Iron Ore IOP changes. Table 13. Autoregressive AR Summary Or R-Squared, AIC And SIC With A Constant, Table 14. Autoregressive AR Summary Of R-Squared, AIC, And SIC Without A Constantspecifically state the values for each information criterion for modelling with a constant and without respectively and can be found in the appendices under Autoregressive Modelling Results Table.

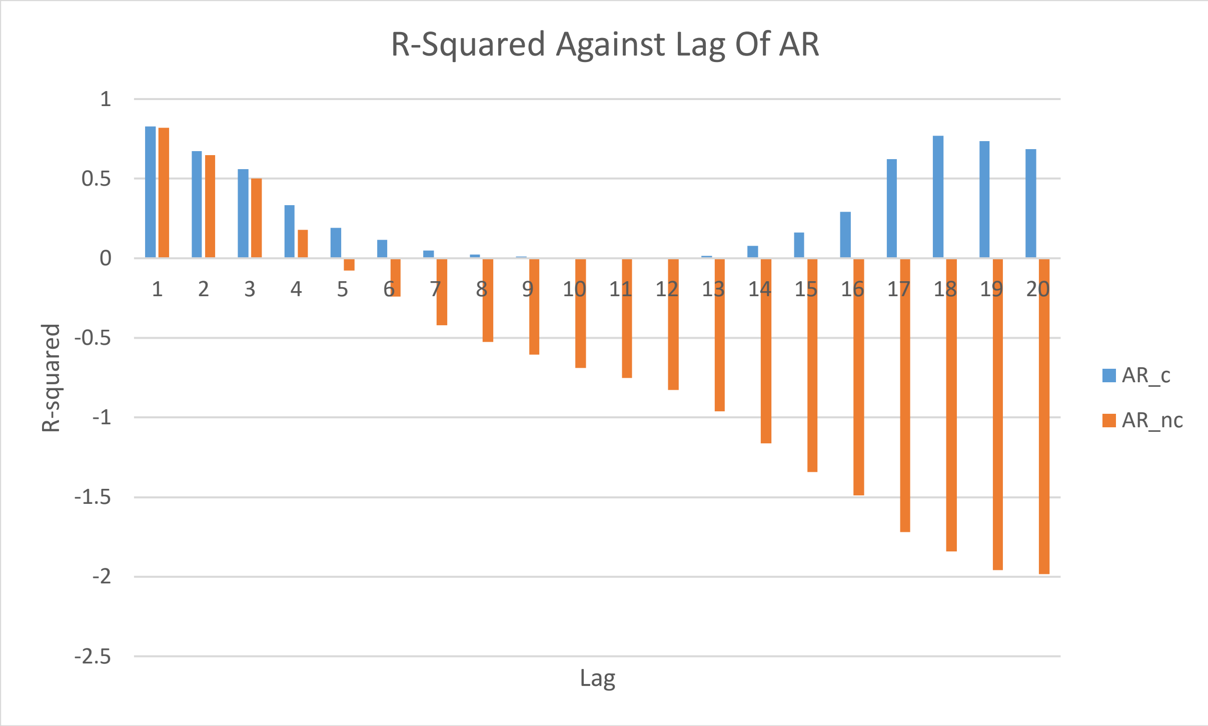

Figure 21. R-Squared Against Lag of AR

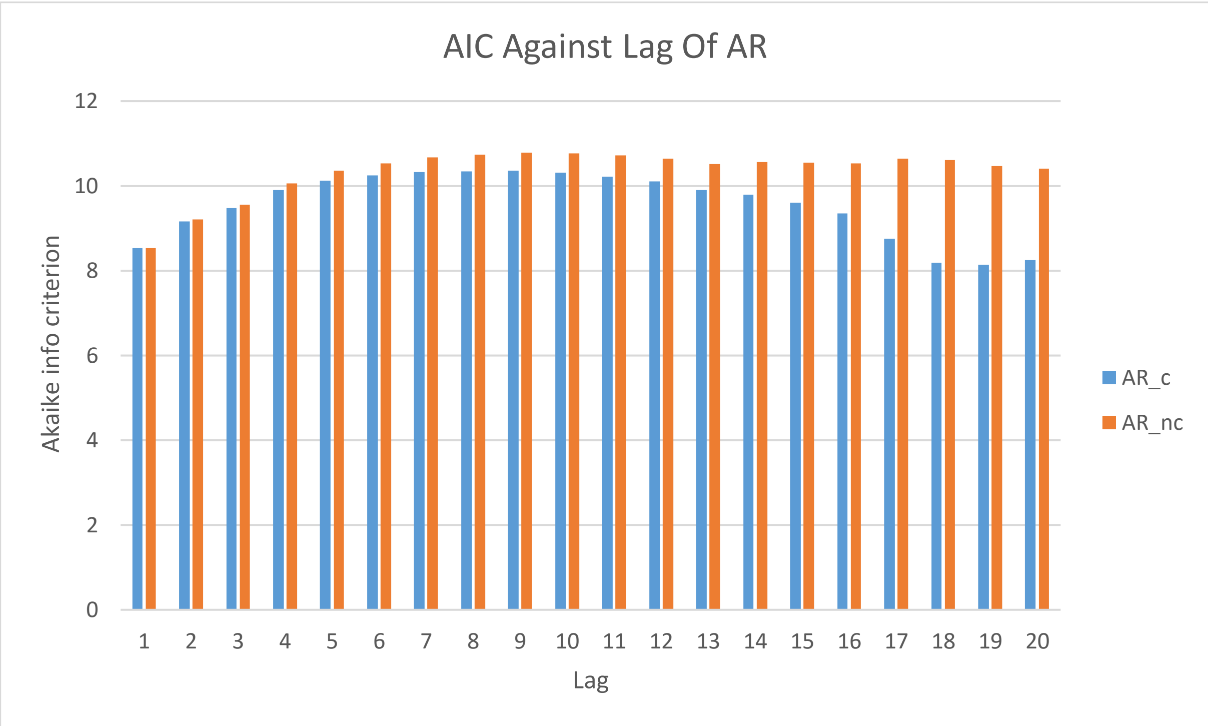

Figure 22. AIC Against Lag Of AR

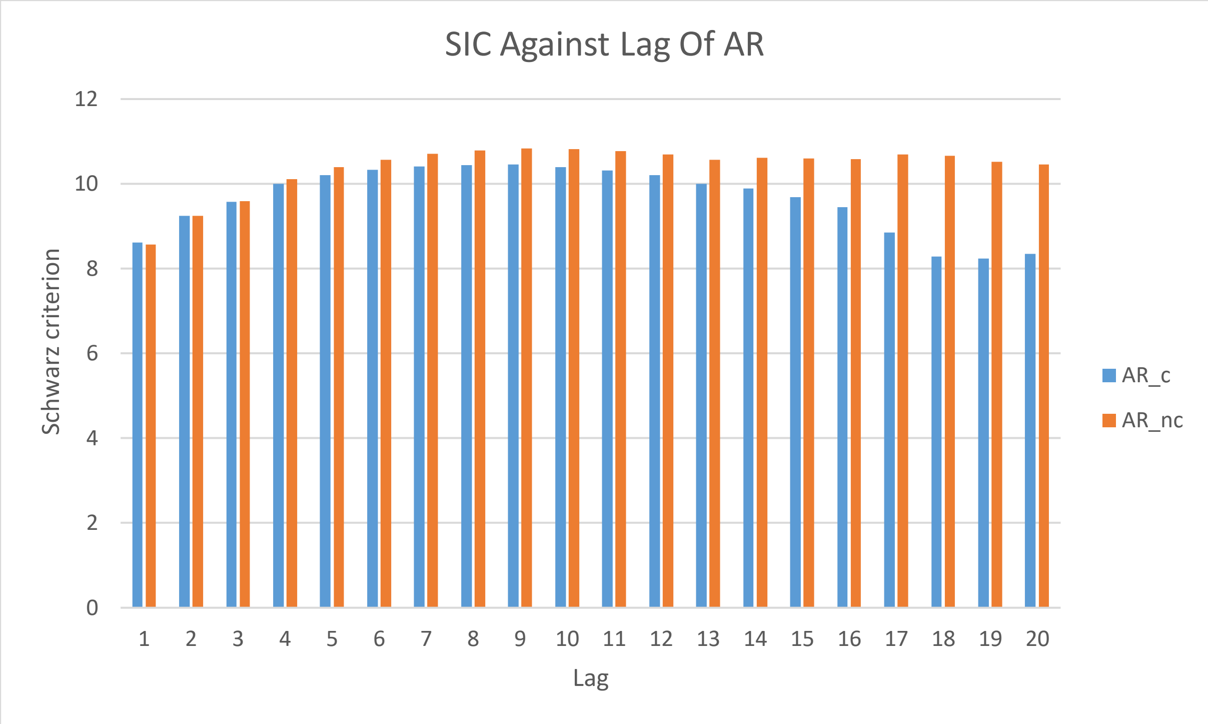

Figure 23. SIC Against Lag Of AR

The highest R^2, both with a constant and without, was observed at a lag of 1 quarter with values of 0.826144 and 0.818106 respectively. The lowest AIC value of reliable observations, both with a constant and without, was observed at a lag of 1 quarter with values of 8.528985 and 8.525406 respectively. The lowest SIC value of reliable observations, both with a constant and without, was observed at a lag of 1 quarter with values of 8.612574 and 8.5672 respectively.

Using the information criteria, it seems that when modelling using the Autoregressive AR form the best lag is of 1 quarter. Since the AIC and SIC values are lower when modelled without a constant compared with when modelled with a constant, it also seems that when modelling using the Autoregressive AR form it is best to use no constant.

The data set used for modelling the price of Iron Ore decreased as the lag of the price of Iron Ore used as an independent variable increased. Values of information criterion were considered less reliable as the lag of the independent variable increased. Specifically, lower values for both AIC and SIC were observed for lags 18, 19 and 20 for modelling with a constant than for lag 1. The reliability of these values is questionable as modelling conducted with 1 lag used 41 observations whereas lag of 17 quarters using 25. It was considered unlikely that lags 17, 18 and 19 were better fits for the data and so they were not considered when determining the best lag variable for the autoregressive.

The best autoregressive model used to model the price of Iron Ore was with a lag of 1 quarter and no constant. Regression modelling determined the respective coefficients with the complete form as follows,

$$IOP_{t-1} = 0.988406\, IOP_{t-1} + e_t. \qquad \text{Equation 23}$$

This relationship was used to infer the values that the model would have returned for the data set used to create it. This is demonstrated in Figure 24. Comparison: IOP and AR1 Estimates.

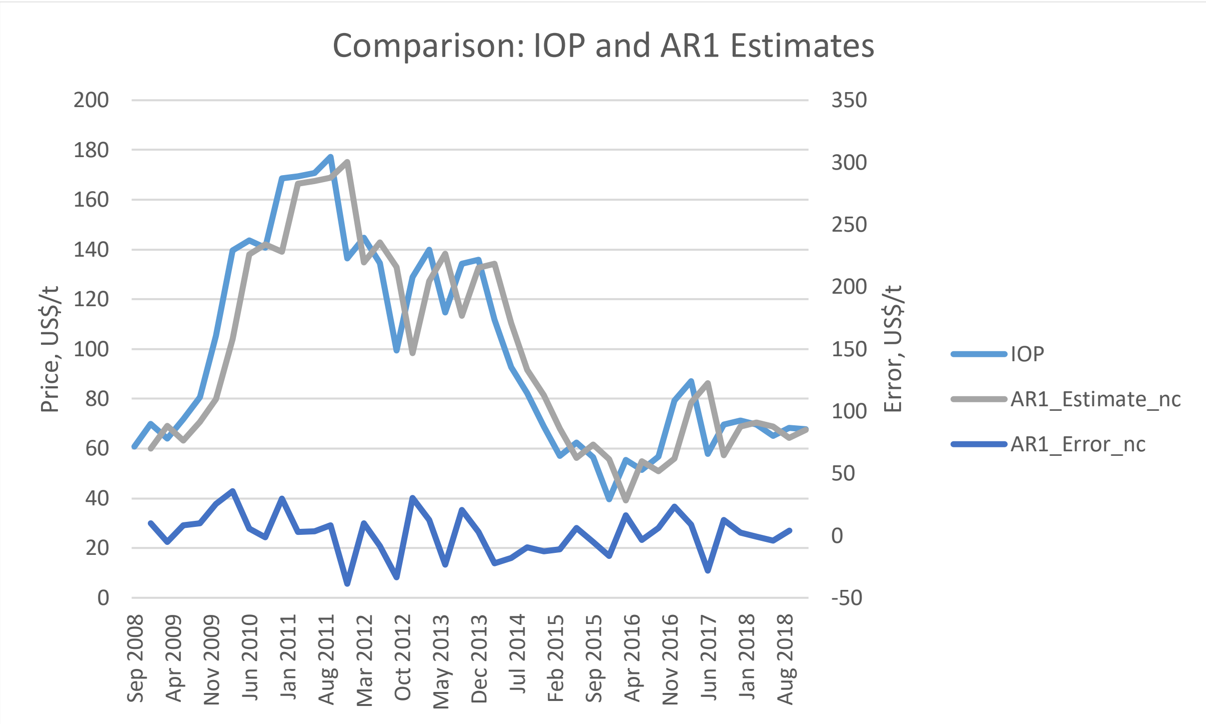

Figure 24. Comparison: IOP and AR1 Estimates

Firstly, with respect to the errors, there seems to be no trend. The error values seem to be fairly evenly distributed through time and are generally reverting to 0.

This Autoregressive model with no constant is, by decision from the information criteria and discretion, the best fit for the price of Iron Ore data and estimates seem quite close to the data set values. This model does not seem to be predicting the price of Iron Ore but rather following it, or being displaced. Unfortunately, as it has a coefficient for the lagged variable of the price of Iron Ore very close to 1 it seems that the model is just copying the previous value from the data set.

9.2.2 LEADING INDICATOR LI

Modelling was conducted using the Leading Indicator LI form, both with a constant and without. The independent variables used for were Total Sales All TSA, Crude Steel Production China CSC, Crude Steel Production World CSW, Real GDP Growth USA GDPU, Real GDP Growth China GDPC, Oil Price West Texas Intermediate OPW, Oil Price Brent Crude OPB, Copper Price CP, and Federal Funds Rate FFR.

The information criterions used to help indicate the Leading Indicator model that fits the data best were R^2, Akaike's Information Criterion AIC and Schwartz Information Criterion SIC. The best lag for each respective independent variable was first found with a summary of all demonstrated in Table 2. Leading Indicator Independent Variable With Constant Highlight specifically stating the values for each information criterion.

Modelling conducted without constants was omitted as they consistently returned negative R^2 values and were so considered unreliability.

Table 2. Leading Indicator Independent Variable With Constant Highlight

Variable	Lag	R^2	AIC	SIC
TSA	3	0.771853	8.863226	8.953012
CSC	9	0.814771	8.685603	8.775389
CSW	8	0.517201	9.644906	9.733783
GDPU	11	0.223952	10.02029	10.10304
GDPC				
OPW	1	0.297217	9.921125	10.00387
OPB	1	0.378013	9.798997	9.881743
CP	1	0.527518	9.524076	9.606822
FFR	20	0.681989	9.128163	9.210909

The Real GDP Growth of China GDPC was omitted from the summary in Table as it did not show any significant lagged variables.

The best independent variable for Leading Indicator modelling with a constant was considered to be Crude Steel Production China CSC with a lag of 9 quarters. This is as it had the highest R^2, lowest AIC and also lowest SIC with values of 0.814771, 8.685683, and 8.775389 respectively. Regression modelling determined the respective coefficients with the complete form as follows,

$$IOP_t = 333.9624 - 1.334368\, CSC_{t-9} + e_t. \qquad \text{Equation 24}$$

This relationship implies that as Crude Steel Production in China CSC increases by 1 Mt in a particular quarter, the price of Iron Ore will fall by US$1.334368 in 9 quarters after considering a constant of 333.9624. This relationship was then used to infer the values that the model would have returned for the data set used to create it. This is demonstrated in Figure 25. Comparison: IOP and LI CSC9 Estimates.

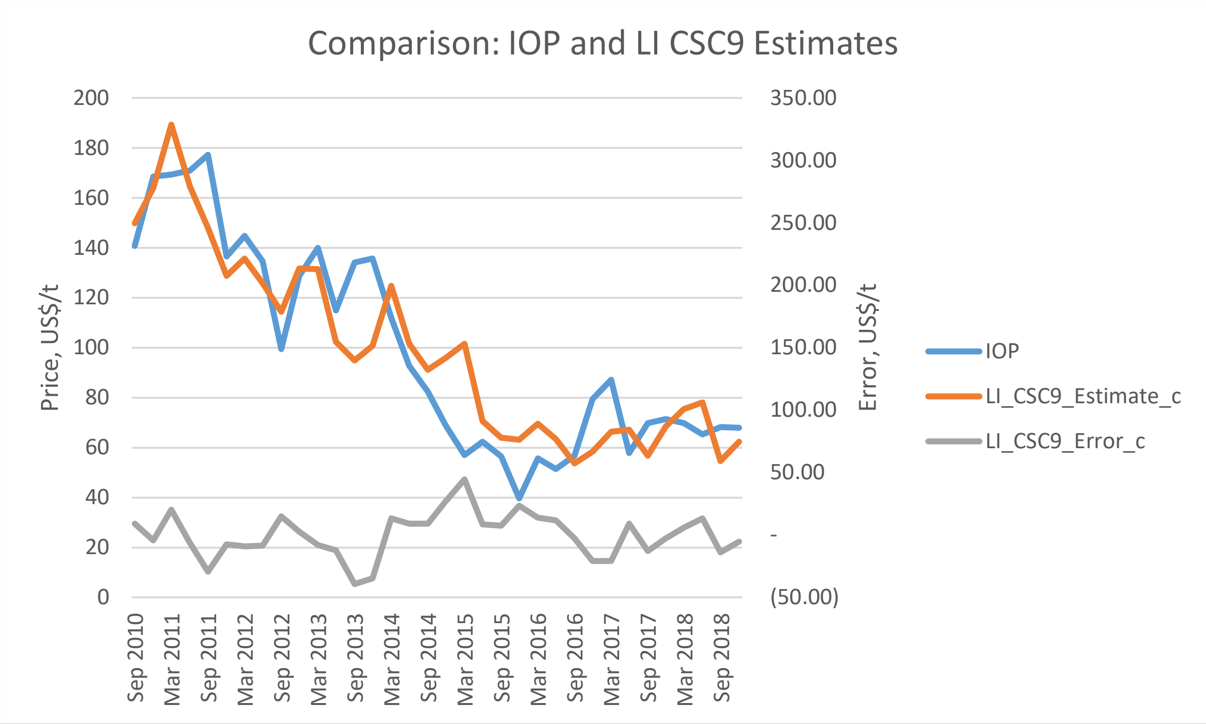

Figure 25. Comparison: IOP and LI CSC9 Estimates

With respect to the errors, there seems to be no trend. The error values seem to be fairly evenly distributed through time and are generally reverting to 0.

This Leading Indicator model using Crude Steel Production in China CSC with a constant is, by decision from the information criteria and discretion, the best fit for the price of Iron Ore data and estimates seem quite close to the data set values. This model does seem to be generally following the trend of the price of Iron Ore rather well although it seems to be lacking definition; it doesn't particularly move with the price data. Overall, the model follows quit quite closely, which is impressive considering it is only using Crude Steel Production in China CSC data to define the relationship, but it does not look accurate enough to predict on its own.

9.2.3 LEADING INDICATOR WITH TWO INDEPENDENT VARIABLES LI2

Modelling was conducted using the Leading Indicator form with two independent variables LI2, both with a constant and without. The independent variables used for modelling were Total Sales All TSA, Crude Steel Production China CSC, Crude Steel Production World CSW, Real

GDP Growth USA GDPU, Real GDP Growth China GDPC, Oil Price West Texas Intermediate OPW, Oil Price Brent Crude OPB, Copper Price CP, and Federal Funds Rate FFR.

The information criterions used to help indicate the Leading Indicator model that fits the data best were R^2, Akaike's Information Criterion AIC and Schwartz Information Criterion SIC. The lag that maximised correlation with the price of Iron Ore for each independent variable was used with all other independent variables. A highlight of models that demonstrated favourable information criterion are shown in Table 3. Leading Indicator With Two Independent Variables and Constant Highlight, specifically stating the independent variables used, their lags and respective values for each information criterion.

Modelling conducted without constants was omitted as they consistently returned less favourable information criterion compared to their respective models with constants.

Table 3. Leading Indicator With Two Independent Variables and Constant Highlight

Variables	Lag	Variable	Lag	R^2	AIC	SIC
TSA	2	CSC	9	0.86143	8.454205	8.588884
TSA	2	CSW	8	0.800935	8.816079	8.949395
TSA	2	GDPU	11	0.81929	8.71934	8.852655
TSA	2	OPW	1	0.805512	8.792816	8.926132
TSA	2	OPB	1	0.803429	8.803472	8.936787
TSA	2	CP	1	0.85613	8.491359	8.624675
TSA	2	FFR	20	0.81328	8.752059	8.885375
CSC	9	CSW	8	0.815633	8.739762	8.874441
CSC	9	GDPU	11	0.846533	8.556319	8.690998
CSC	9	GDPC	1	0.800404	8.676916	8.815689
CSC	9	OPW	1	0.835184	8.627662	8.762341
CSC	9	OPB	1	0.8391	8.603617	8.738296
CSC	9	CP	1	0.835366	8.62656	8.761239
CSC	9	FFR	20	0.835815	8.623826	8.758505

The best model for Leading Indicator with two independent variables and a constant was considered to be with Total Sales All TSA with a lag of 2 quarters and Crude Steel Production China CSC with a lag of 9 quarters. This is as it had the highest R^2, lowest AIC and also lowest SIC with values of 0.86143, 8.454205, and 8.588884 respectively. Regression modelling determined the respective coefficients with the complete form as follows,

$$IOP_{\mathrm{t}} = 329.1396 - 0.39577\, TSA_{t-2} - 0.800485\, CSC_{t-9} + e_t. \qquad \text{Equation 25}$$

This relationship implies that as Total Sales All TSA increases by 1 Mt in a particular quarter, the price of Iron Ore will fall by US$0.39577 in 2 quarters and that as Crude Steel Production in China CSC increases by 1 Mt in a particular quarter, the price of Iron Ore will fall by US$1.334368 in 9 quarters. This is after considering a constant of 329.1396. This relationship was then used to infer the values that the model would have returned for the data set used to create it. This is demonstrated in Figure 26. Comparison: IOP and LI2 Estimates.

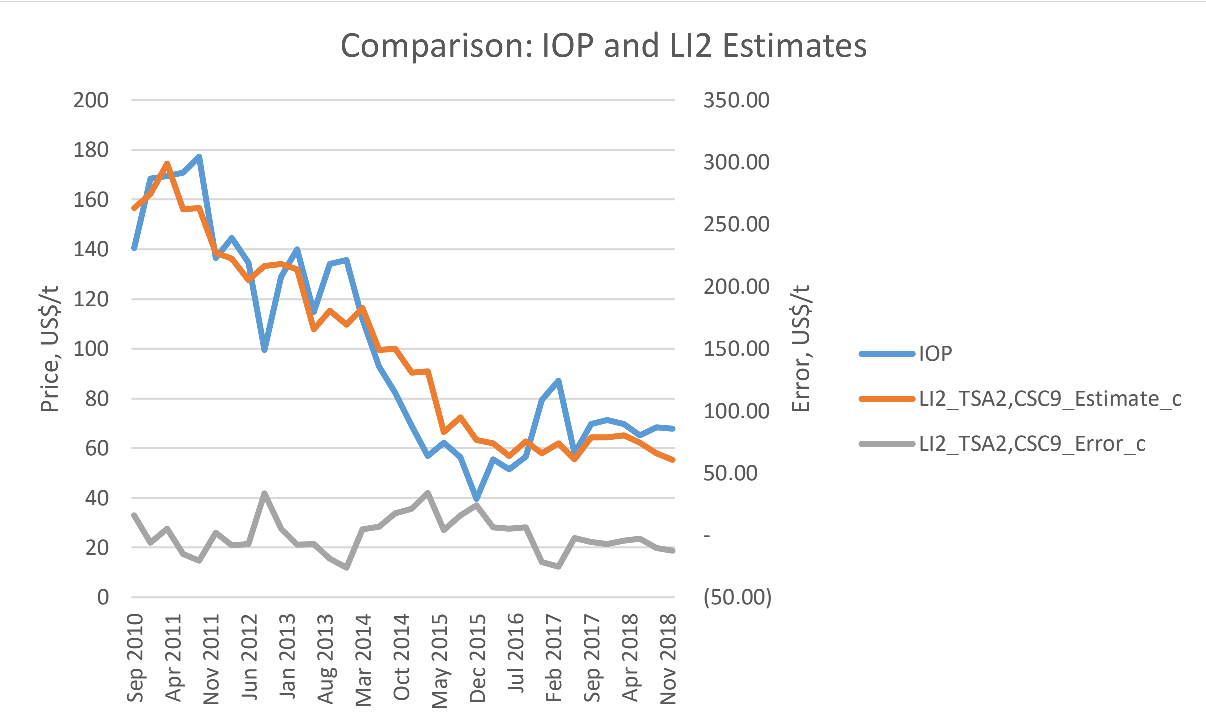

Figure 26. Comparison: IOP and LI2 Estimates

With respect to the errors, there seems to be no trend. The error values seem to be fairly evenly distributed through time and are generally reverting to 0.

The Leading Indicator with Total Sales All TSA and Crude Steel Production in China CSC as independent variables with a constant is, by decision from the information criteria and discretion, the best fit for the price of Iron Ore data and estimates seem quite close to the data set values. This model does follow the trend of the price of Iron Ore very well although there are instances where movements in the estimate are not of the same magnitude as the data and other instances where it seems that they move in opposite directions. Overall, the model follows quit quite closely, which is impressive considering it is only using Total Sales All and Crude Steel Production in China CSC data to define the relationship, but it does not look accurate enough to predict on its own.

9.2.4 AUTOREGRESSIVE DISTRIBUTED LAG ARDL

Modelling was then conducted using the Autoregressive Distributed Lag form, both with a constant and without. The independent variables used for modelling were the price of Iron Ore with a lag of 1 quarter and the respective lag maximising correlation with the price of Iron Ore for Total Sales All TSA, Crude Steel Production China CSC, Crude Steel Production World CSW, Real GDP Growth USA GDPU, Real GDP Growth China GDPC, Oil Price West Texas Intermediate OPW, Oil Price Brent Crude OPB, Copper Price CP, and Federal Funds Rate FFR.

The information criterions used to help indicate the Autoregressive Distributed Lag model that fits the data best were R^2, Akaike's Information Criterion AIC and Schwartz Information Criterion SIC. The lag that maximised correlation with the price of Iron Ore for each independent variable was used with all other independent variables. A highlight of models that demonstrated favourable information criterion are shown in Table 4. Autoregressive Distributed Lag With Constant Summary, specifically stating the independent variables used, their lags and respective values for each information criterion.

Modelling conducted without constants was omitted as they consistently returned less favourable information criterion compared to their respective models with constants.

Table 4. Autoregressive Distributed Lag With Constant Summary

Variables	Lag	Variable	Lag	R^2	AIC	SIC
IOP	1	TSA	2	0.875131	8.349712	8.483028
IOP	1	CSC	9	0.878149	8.325634	8.460313
IOP	1	CSW	8	0.860486	8.460615	8.59393
IOP	1	GDPU	11	0.836408	8.516916	8.6423
IOP	1	GDPC	1	0.841838	8.444238	8.583011
IOP	1	OPW	1	0.826643	8.574896	8.700279
IOP	1	OPB	1	0.826269	8.577048	8.702431
IOP	1	CP	1	0.830317	8.553473	8.678857
IOP	1	FFR	20	0.850567	8.426389	8.551773

The best model for Autoregressive Distributed Lag with a constant was considered to be with the price of Iron Ore IOP with a lag of 1 quarter and Crude Steel Production China CSC with a lag of 9 quarters. This is as it had the highest R^2, lowest AIC and also lowest SIC with values of 0.878149, 8.325634, and 8.460313 respectively. Regression modelling determined the respective coefficients with the complete form as follows,

$$IOP_t = 151.0296 - 0.546875\, IOP_{t-1} - 0.609997\, CSC_{t-9} + e_t. \qquad \text{Equation 26}$$

This relationship implies that as Total Sales All TSA increases by 1 Mt in a particular quarter, the price of Iron Ore will fall by US$0.546875 in 1 quarters and that as Crude Steel Production in China CSC increases by 1 Mt in a particular quarter, the price of Iron Ore will fall by US$0.609997 in 9 quarters. This is after considering a constant of 151.0296. This relationship was then used to infer the values that the model would have returned for the data set used to create it. This is demonstrated in Figure 27. Comparison: IOP and ARDL Estimates.

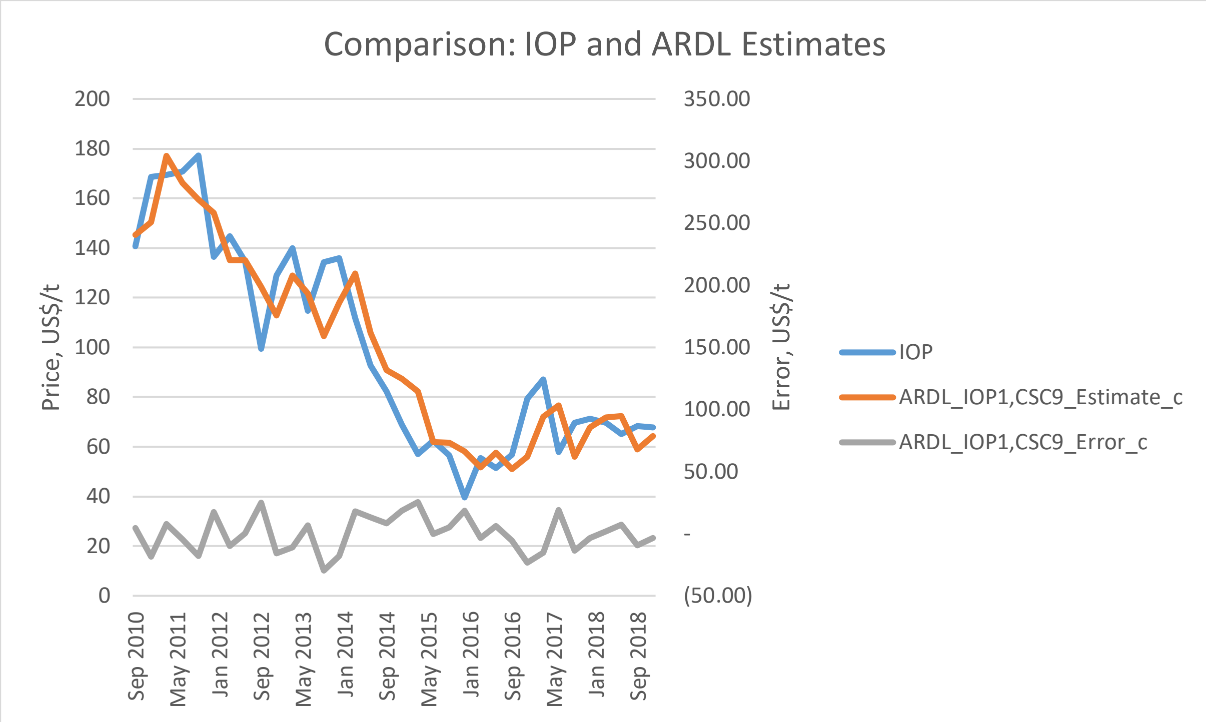

Figure 27. Comparison: IOP and ARDL Estimates

With respect to the errors, there seems to be no trend. The error values seem to be fairly evenly distributed through time and are generally reverting to 0.

The Autoregressive Distributed Lag ARDL using a lag of 1 quarter for the price of Iron Ore IOP, using a lag of 9 quarters for Crude Steel Production in China CSC as independent variables with a constant is, by decision from the information criteria and discretion, the best fit for the price of Iron Ore data and estimates seem quite close to the data set values. This model does follow the trend of the price of Iron Ore very well although it seems towards the middle and end of the data set that the model is lagging behind the data, almost seeming displaced. There are also instances where movements in the estimate are not of the same magnitude as the data and other instances where it seems that they move in opposite directions. Overall, the model follows quit quite closely but it does not look accurate enough to predict on its own.

9.2.5 AUTOREGRESSIVE MOVING AVERAGE ARMA

Modelling was conducted using the Autoregressive Moving Average form both with a constant and without. The independent variables used for modelling were lagged values of the price of

Iron Ore and lagged values of errors from autoregressive modelling, found in Table 15. Error Values From Autoregressive Modelling in the appendices Errors From Autoregressive AR1 Modelling. All combinations were tested using up to 4 lags for both independent variables.

The information criterions used to help indicate ARMA models that fit the data best were R^2, Akaike's Information Criterion AIC and Schwartz Information Criterion SIC. A highlight of models that demonstrated favourable information criterion are shown in Table 5. ARMA With Constant Highlight, specifically stating the independent variables used, their lags and respective values for each information criterion.

Modelling conducted without constants was omitted as they consistently returned less favourable information criterion compared to their respective models with constants.

Table 5. ARMA With Constant Highlight

Variables	Lag	Variable	Lag	R^2	AIC	SIC
IOP	1	Error	1	0.824347	8.603096	8.729762
IOP	1	Error	2	0.822383	8.622827	8.750793
IOP	1	Error	3	0.873998	8.295791	8.425074
IOP	1	Error	4	0.820617	8.67298	8.803595
IOP	2	Error	1	0.824347	8.603096	8.729762
IOP	2	Error	2	0.668761	9.24604	9.374006
IOP	2	Error	3	0.710534	9.127535	9.256818
IOP	2	Error	4	0.695157	9.203251	9.333866
IOP	3	Error	1	0.691266	9.17568	9.303646
IOP	3	Error	2	0.668761	9.24604	9.374006
IOP	3	Error	3	0.600591	9.449484	9.578767
IOP	3	Error	4	0.574722	9.536197	9.666812
IOP	4	Error	1	0.427902	9.808807	9.93809
IOP	4	Error	2	0.418563	9.824999	9.954282
IOP	4	Error	3	0.600591	9.449484	9.578767
IOP	4	Error	4	0.364252	9.938257	10.06887

The best model for Autoregressive Moving Average ARMA modelling with a constant was considered to be with a lag of 1 quarter for the price of Iron Ore and a lag of 3 quarters for the error term. This is as it had the highest R^2, lowest AIC and also lowest SIC with values of 0.873998, 8.295791, and 8.425074 respectively. Regression modelling determined the respective coefficients with the complete form as follows,

$$IOP_t = 12.809 - 0.870972\, IOP_{t-1} - 0.539103\, u_{t-3} + e_t. \qquad \text{Equation 27}$$

This relationship implies that as the price of Iron Ore IOP increases by $1 in a particular quarter, the price of Iron Ore will decrease by US$0.87092 in the next quarter and that as the error term increases by $1 in a particular quarter, the price of Iron Ore will decrease by US$0.5391031 in 3 quarters. This is after considering a constant of 12.809. This relationship was then used to infer the values that the model would have returned for the data set used to create it. This is demonstrated in Figure 28. Comparison: IOP and ARMA With Constant.

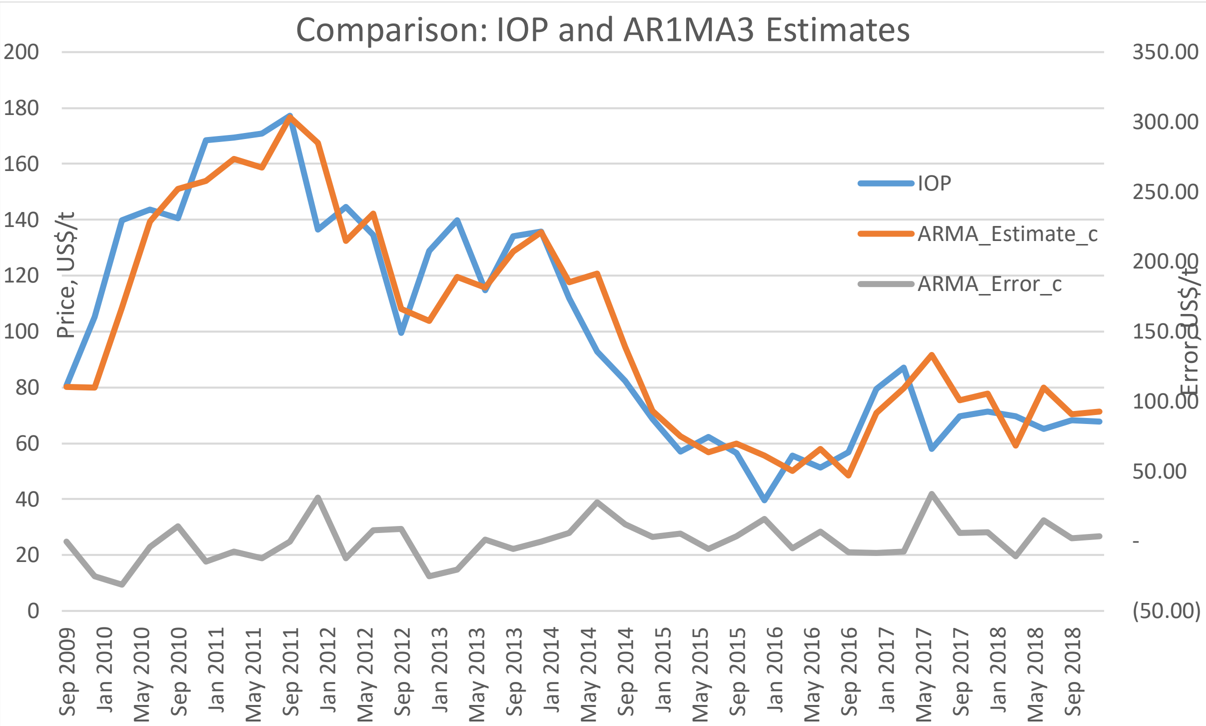

Figure 28. Comparison: IOP and ARMA With Constant

With respect to the errors, there seems to be no trend. The error values seem to be fairly evenly distributed through time although are generally positive in the second half of the data set.

This Autoregressive Moving Average ARMA model with a constant is, by decision from the information criteria and discretion, the best fit for the price of Iron Ore data and estimates seem quite close to the data set values. This model does follow the trend of the price of Iron Ore very well although there are instances where movements in the estimate are not of the same magnitude as the data and other instances where it seems that they move in opposite directions evident towards the end of the data set. Overall, the model follows quit quite closely and is generally quite accurate although it may need attention and improvements in certain areas to increase accuracy.

9.2.6 AUTOREGRESSIVE MOVING AVERAGE WITH LEADING INDICATOR ARMAX

Modelling was then conducted using the Autoregressive Moving Average form with an added leading indicator, both with a constant and without. The independent variables used for modelling were the price of Iron Ore with a lag of 1 quarter, the values of errors terms with a lag of 3 quarters from autoregressive modelling and a lagged independent variable maximising correlation with the price of Iron Ore from Total Sales All TSA, Crude Steel Production China CSC, Crude Steel Production World CSW, Real GDP Growth USA, Real GDP Growth China, Oil Price West Texas Intermediate, Oil Price Brent Crude OPB, Copper Price CP and Federal Funds Rate FFR. The values used for error term values can be found in Table 15. Error Values From Autoregressive Modelling in the appendices Errors From Autoregressive AR1 Modelling.

The information criterions used to help indicate ARMA models that fit the data best were R^2, Akaike's Information Criterion AIC and Schwartz Information Criterion SIC. Models that demonstrated favourable information criterion are shown in Table 6. ARMA With Leading Indicator And Constant, specifically stating the independent variables used, their lags and respective values for each information criterion.

Modelling conducted without constants was omitted as they consistently returned less favourable information criterion compared to their respective models with constants.

Table 6. ARMA With Leading Indicator And Constant

Variable	Lag	Variable	Lag	Variable	Lag	R^2	AIC	SIC
IOP	1	Error	3	TSA	2	0.91538	8.017762	8.195516
IOP	1	Error	3	CSC	9	0.922061	7.937589	8.117161
IOP	1	Error	3	CSW	8	0.909223	8.087996	8.26575
IOP	1	Error	3	GDPU	11	0.874224	8.346633	8.519011
IOP	1	Error	3	GDPC	1	0.88999	8.145706	8.330736
IOP	1	Error	3	OPW	1	0.874488	8.344527	8.516904
IOP	1	Error	3	OPB	1	0.874245	8.346467	8.518844
IOP	1	Error	3	CP	1	0.883413	8.270763	8.443141
IOP	1	Error	3	FFR	20	0.892871	8.186162	8.358539

The best model for Autoregressive Moving Average ARMA modelling with a leading indicator and a constant was considered to be with a lag of 1 quarter for the price of Iron Ore, a lag of 3 quarters for the error term and a lag of 9 quarters for Crude Steel Production in China CSC. This is as it had the highest R^2, lowest AIC and also lowest SIC with values of 0.922061, 7.937589, and 8.117161 respectively. Regression modelling determined the respective coefficients with the complete form as follows,

$$IOP_t = 134.0886 + 0.566401\, IOP_{t-1} + 0.477979\, u_{t-3} - 0.524117\, CSC_{t-9} + e_t.$$

Equation 28

This relationship implies that as the price of Iron Ore IOP increases by $1 in a particular quarter the price of Iron Ore IOP will increase by US$0.566401 in the next quarter, as the error term increases by $1 in a particular quarter the price of Iron Ore will increase by US$0.477979 in 3 quarters and as Crude Steel Production in China increases by 1 Mt in a particular quarter the price of Iron Ore will fall by US$0.524117. This is after considering a constant of 134.0886. This relationship was then used to infer the values that the model would have returned for the data set used to create it. This is demonstrated in Figure 29. Comparison: IOP and ARMA With CSC9 and Constant.

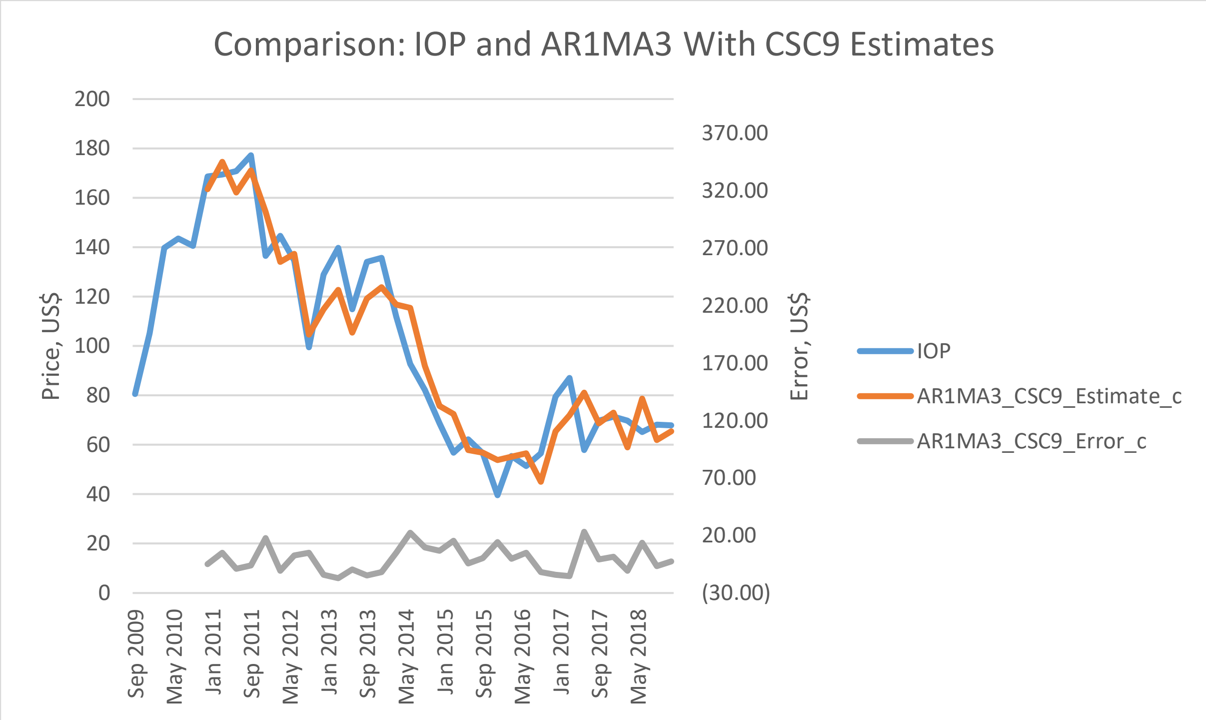

Figure 29. Comparison: IOP and ARMA With CSC9 and Constant

With respect to the errors, there seems to be no trend. The error values seem to be fairly evenly distributed through time although are generally positive in the second half of the data set.

This Autoregressive Moving Average ARMA model with Crude Steel Production in China CSC as a variable and a constant is, by decision from the information criteria and discretion, the best fit for the price of Iron Ore data and estimates seem quite close to the data set values. This model does follow the trend of the price of Iron Ore very well and interestingly it is often picking the right inflection points. This is particularly evident between March 2012 and March 2015 where the model is consecutively picking the right times to rise and fall with the model. It does, however, fail to get the magnitude of these rises and falls right. Overall, the model follows quit quite closely and generally seems quite accurate.

9.3 SUMMARY OF MODELS

9.3.1 BEST OF THE BEST

A summary of the best variations for each model form are demonstrated in Table 7. Summary Of Best Models For Each Form. These models include the Autoregressive AR, Leading Indicator LI, Leading Indicator with two independent variables LI2, Autoregressive Distributed Lag ARDL, Autoregressive Moving Average ARMA and the Autoregressive Moving Average with an independent variable ARMAX. All models in this summary are used with a constant except for the Autoregressive AR model.

Table 7. Summary Of Best Models For Each Form

Model	Variable	Lag	Variable	Lag	Variable	Lag	R^2	AIC	SIC
AR	IOP	1					0.818106	8.525406	8.5672
LI	CSC	9					0.814771	8.685603	8.775389
LI2	TSA	2	CSC	9			0.86143	8.454205	8.588884
ARDL	IOP	1	CSC	9			0.878149	8.325634	8.460313
ARMA	IOP	1	Error Term	3			0.873998	8.295791	8.425074
ARMAX	IOP	1	Error Term	3	CSC	9	0.922061	7.937589	8.117161

The overall most suitable model to fit the price of Iron Ore data, used in this investigation, was deemed to be Autoregressive Moving Average model with a leading indicator variable and constant. This model used the price of Iron Ore IOP with a lag of 1 quarter, the error term with a lag of 3 quarters, Crude Steel Production China CSC with a lag of 9 quarters and a constant. This is this most suitable model as it has the highest as it has the R^2 value, lowest AIC value and lowest SIC value with 0.922061, 7.937589 and 8.117161 respectively.

Therefore, this Autoregressive Moving Average with a leading indicator ARMAX model is the most suitable form for forecasting.

9.3.2 FORECASTING

The Autoregressive Moving Average model with a leading indicator ARMAX was determined to explain the price of Iron Ore IOP data best and will therefore be used to forecast.

The data used for regression modelling in this investigation has a range from September 2009 to December 2018. Forecasting will be conducted for the price of Iron Ore in March 2019. A summary of the forecast with comparison to the real price of Iron Ore for March 2019 can be found in Table 8. March 2019 Forecast Summary.

Table 8. March 2019 Forecast Summary

Date	IOP, US$/t	Model Forecast, US$/t	Model Error, US$/t	Model Error, %
March 2019	85.75	64.42	(21.33)	(0.25)

The Autoregressive Moving Average model with a leading indicator estimated that the price of Iron Ore in March 2019 would be US$64.42/t. This is considering the model is estimating using lagged values of the price of Iron Ore IOP, the error term and Crude Steel Production in China CSC. In reality, the price of Iron Ore for March 2019 was US$85.75/t which represents a model error of negative US$21.33/t and approximately negative 25% difference.

This is a significant difference between the model and reality which could likely be explained by shocks from external market factors beyond the scope of this investigation.

9.3.3 THE QUILT

This investigation is able to stitch together the best performing models at various lags which strictly have Akaike’s Information Criterion AIC and Schwartz Information Criterion SIC values below 9. This will enable suggestion of which model may be best to use depending on the amount of time in the future concerned.

The models used to create The Quilt are demonstrated in Table 9. The Quilt: Summary Of Models At Each Lag. The models used were the ARMAX model at a lag of 1 quarter, the Leading Indicator with two independent variables at a lag of 2 quarters, the Leading Indicator model at a lag of 3 quarters and another Leading Indicator model at a lag of 9 quarters. It must be noted that they all use a constant.

Table 9. The Quilt: Summary Of Models At Each Lag

	Variable	Lag	Variable	Lag	Variable	Lag
ARMAX	IOP	1	Error Term	3	CSC	9
Leading Indicator LI2	TSA	2	CSC	9		
Leading Indicator LI2	CSC	9	GDPU	11		
Leading Indicator LI	TSA	3				

When modelling at a lag of 1 quarter the ARMAX model is best. It uses the price of Iron Ore with a lag of 1 quarter, the error term with a lag of 3 quarters and Crude Steel Production in China CSC with a lag of 9 quarters as well as a constant. The formula is denoted as follows,

$$IOP_t = 134.0886 + 0.566401\, IOP_{t-1} + 0.477979\, u_{t-3} - 0.524117\, CSC_{t-9} + e_t.$$

Equation 29

When modelling at a lag of 2 quarters the Leading Indicator LI2 model is best. It uses Total Sales All TSA with a lag of 2 quarters, and Crude Steel Production in China CSC with a lag of 9 quarters as well as a constant. The formula is denoted as follows,

$$IOP_t = 329.1396 - 0.39557\, TSA_{t-2} - 0.800485\, CSC_{t-9} + e_t.$$ Equation 30

When modelling at a lag of 3 quarters the Leading Indicator LI model is best. It uses Total Sales All TSA with a lag of 3 quarters as well as a constant. The formula is denoted as follows,

$$IOP_t = 283.1059 - 0.829038\, TSA_{t-3} + e_t.$$ Equation 31

When modelling at a lag of 9 quarters the Leading Indicator LI2 model is best. It uses Crude Steel Production in China with a lag of 9 quarters, and Real GDP Growth USA with a lag of 11 quarters as well as a constant. The formula is denoted as follows,

$$IOP_t = 318.7476 - 1.226793\, CSC_{t-9} - 11.34577\, GDPU_{t-11} + e_t.$$ Equation 32

The calculations for 95% prediction interval spreads, at each respective lag, are demonstrated in Table 10. The Quilt: 95% Prediction Interval Spread Summary. This will enable a forecast confidence interval spread, demonstrated in Figure 30. The Quilt: 95% Prediction Interval Spread.

Table 10. The Quilt: 95% Prediction Interval Spread Summary

Lag	1	2	3	4
Standard Error	12.11919	15.89676	19.76996	16.72948
P-Value	0.025	0.025	0.025	0.025
Number of Observations	34	34	34	34
Corresponding t-Value	2.042	2.042	2.042	2.042
Prediction Spread, +/- US$/t	12.11919	15.89676	19.76996	16.72948

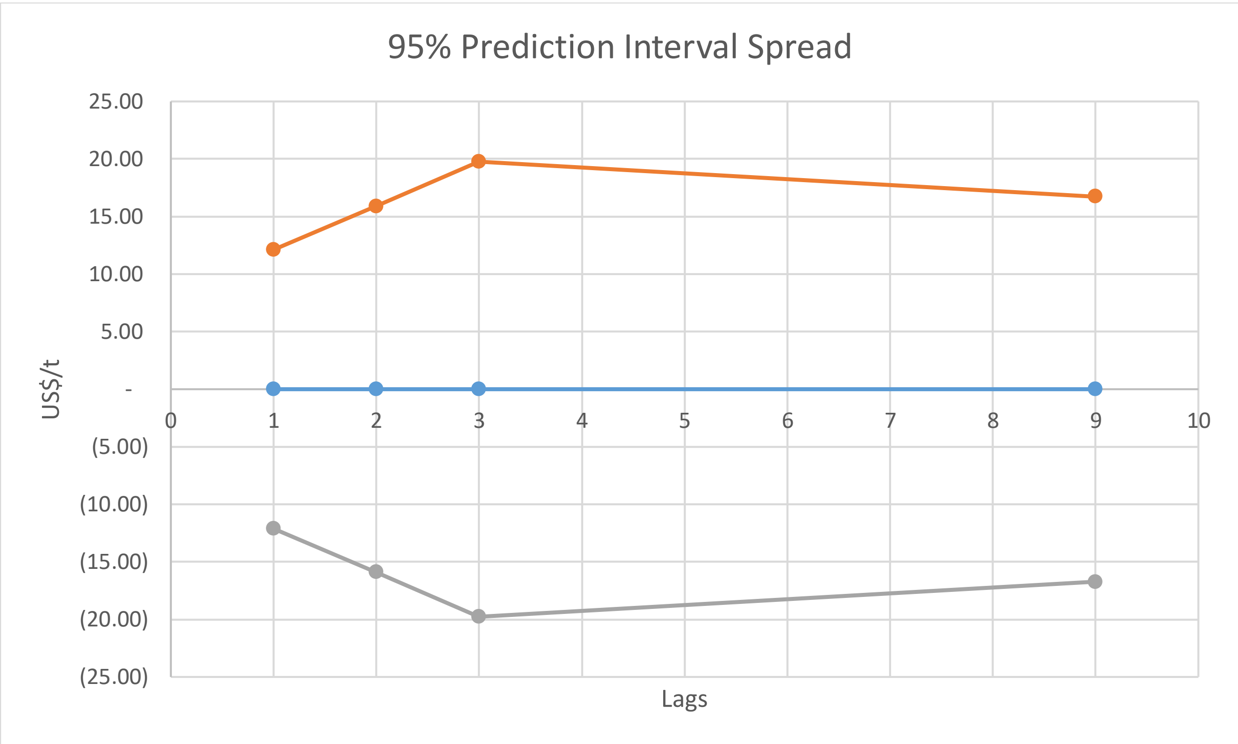

Figure 30. The Quilt: 95% Prediction Interval Spread

As demonstrated in Figure 30. The Quilt: 95% Prediction Interval Spread, the 95% prediction interval spread increases from lags of 1 quarter to 3. As there is no significant data for lags of 4 quarters to 8 it is presumed that the prediction interval spread decreases by increasing magnitude to lag 9.

10. DISCUSSION

The first analysis conducted was correlation with the price of Iron Ore, investigating at which lag it was maximised with each independent variable. There was no minimum correlation level set to include or exclude variables as it was not expected that high correlations would be observed. This is because it was assumed that the price of Iron Ore would likely be affected by several factors.

It was expected that most independent variables would have optimum lags at 1 quarter. This was demonstrated with the price of Iron Ore with 0.91, Real GDP Growth China with 0.65, Oil Price WTI with 0.55, Oil Price Brent Crude with 0.62, and Copper Price with 0.73. It was unexpected for Real GDP Growth China to have such a small lag which may either be a revelation or a discrepancy with the retrieved data.

Independent variables that demonstrated longer optimum lags were Total Sales All with, Crude Steel Production in China with a lag of 2 quarters and correlation of -0.89, Crude Steel Production in World with a lag of 9 quarters and correlation of -0.9, Real GDP Growth USA with a lag of 11 quarters and correlation of -0.473, and Federal Funds Rate with a lag of 20 quarters and correlation of 0.83. These results were all unexpected; particularly such a long optimal lag for Federal Funds Rate.

The very existence of these lags, other than a lag of 1 quarter, with such high correlations with the price of Iron Ore is intriguing in itself but unfortunately beyond the scope of this investigation. Therefore, the cause of these lags occurring when they do will not be investigated.

It was assumed that the lag that, for each independent variable used, maximised correlation with the price of Iron Ore would be the best lagged variables to include when conducting modelling. This assumption allowed the attention of the investigation to be more focused as, without it, the number of equations as well as time needed would have increased significantly.

In reality, this assumption worked well but when modelling using only one independent variable the lag from 1 quarter to 20 quarters was modelled it was noticed that it did not always hold. An example of this is demonstrated when conducting Leading Indicator LI modelling; the optimum lag of Total Sales All TSA is 2 quarters but the optimum model when using this independent variable actually used a lag of 3 quarters. This allows questioning of the models used later in this investigation, with more than one independent variable used, when only the lagged variables maximising correlation with the price of Iron Ore were used. To do more exhaustive modelling with absolutely every lag of every independent variable would increase the number of models and time used incredibly and would likely not be the best use of time.

The models used in this investigation were the Autoregressive AR, Leading Indicator LI, Leading Indicator with two independent variables LI2, Autoregressive Distributed Lag ARDL, Autoregressive Moving Average ARMA and Autoregressive Moving Average with a leading indicator ARMAX.

The ARMA model with a leading indicator, using the price of Iron Ore with a lag of 1 quarter, the error term with a lag of 3 quarters and Crude Steel Production in China with a lag of 9 quarters was considered the best model since it had the highest R^2, lowest AIC value and the lowest SIC value of 0.922061, 7.937589 and 8.117161 respectively.

This model seemed to fit the general trend of the price of Iron Ore data quite well. Interestingly further than that, it generally picked the right times to inflect; it would move in the same direction as the data. This is evident from June 2011 to December 2011, June 2012 to March 2015, June 2015 to March 2016, and September 2016 to March 2017. This likely demonstrates that the independent variables used to model the price of the Iron Ore are suitable and that the model isn't missing any tangible variables.

On the occasions that the model did not move in the same direction as the data, it would generally try to compensate in the following quarter which is likely due to the presence in the model of the lagged error term. This is evident from March 2012 to June 2012, June 2016 to September 2016 and March 2017 to December 2018.

This model did not, however, meet the magnitude of these inflections often. This likely demonstrates that this model fails to quantify some phenomena which is likely related to a mismatch in supply and demand. When supply, of either Crude Steel Production in China or Total Sales All, is lower than demand the cause would be an increase in the price of Iron Ore. This would explain the times when the model inflects at the right point but does not rise as high as the data. Alternatively, it would also explain the times demand is lower than supply which would cause the price of Iron Ore to fall. This would explain the times when the model inflects at the right point but does not fall as low as the data. These occurrences are evident from June 2012 to March 2015.

This could be addressed with an extra independent variable demonstrating the mismatch between supply and demand, of either Iron Ore or Crude Steel, depending on the particular model. The issue with this is that demand would need to be modelled itself which would drift this investigation closer towards the realms of economic reasoning. Not only would it be difficult but it would also mean that modelling of the price of Iron Ore would be using circumstantial values as well as tangible ones, such as Total Sales All TSA or Crude steel Produced in China CSC.

A major assumption in this investigation was that the relationships between independent variables and the price of Iron Ore, denoted by the coefficients from modelling, would hold when forecasting.

Forecasting was conducted using the Autoregressive Moving Average ARMA model with the leading indicator variable; price of Iron ore with a lag of 1 quarter, error term with 3 and Crude Steel Production in China with 9. This model used data up to December 2018 and was used to forecast the price of Iron Ore in March 2019.

This ARMA model predicted that the price of Iron Ore would be US$64.42/t in March 2019. In reality, the price of Iron Ore was actually US$85.75/t, representing an error of negative US$21.33/t and approximately negative 25%. The difference between the model estimate and reality was quite large and demonstrates the weakness of not only this model, or even any of the other models in this investigation; predictive modelling can't anticipate shocks to the market and so generally work best in more stable and stationary times.

The cause of this shock is likely the failure of a tailings dam at the Brumadinho Iron Ore mine in Brazil, owned by Vale, in mid-January 2019. This mine was considered a significant producer, not only for Vale, but for the market. Production at Brumadinho immediately stopped, representing a shock loss of supply to the market. The Iron Ore price immediately began to climb after the incident, continuing to rise in the weeks following with no signs of slowing down; at the time of writing the price of Iron Ore is US$94.38/t.

This climb in the price of Iron Ore likely represents, not only a shock loss of supply to the market, but also speculation as interested parties investigate how long this supply loss will be sustained for.

Admittedly the most suitable model used to forecast the price of Iron Ore does not consider Iron Ore supply, denoted in this investigation as Total Sales All TSA. Even if it had though, it would not have been able to foresee this event occurring due to the frequency of periods; this investigation uses quarterly data and Total Sales All TSA data from previous periods would not have pointed to this event occurring.

There are more variables that cause movements in the price of Iron Ore than are modelled in this investigation, particularly in the short term. The models in this investigation used tangible independent variables to estimate relationships with the price of Iron Ore and this intuition is not believed to be flawed although the reality is that events and speculation beyond the scope of this investigation caused real shocks to the price of Iron Ore. It is unfortunate that the shock loss of supply occurred as the opportunity to compare the ARMA model to ordinary times is now lost. It may now be that the relationships between the independent variables used in this investigation to model the price of Iron Ore will change significantly as their relationships are dynamic.

This investigation used all data available to model the price of Iron Ore, from September 2009 to December 2018. This allowed only one observation to compare predictions from the model with reality at March 2019. In hindsight, it may have been more insightful to reduce the observations used to create the model to allow more opportunities for comparison. This was not conducted initially as it was believed to be in the best interest of this investigation to give as much data to the model as possible in the hope that it would enable more robust relationships to be quantified. Again, considering such an unlikely event to tangibly change the market for Iron Ore has occurred in the same quarter as this investigation was to compare predictions from the model it is recommended that further investigations give model testing more opportunities.

The best model in this investigation, the ARMAX, uses various past independent variables to enable prediction 1 quarter into the future. In reality, this amount of time is likely insignificant to enable strategic mine planning decisions. The ability of mine planning to implement strategies increases as a model increases the time ahead it is able to accurately forecast although this becomes more difficult as time increases.

The purpose of this investigation was to determine the best way to explain the relationship between independent variables and the price of Iron Ore, regardless of the lags involved. The best model from this investigation was an ARMA model that needs data from as recent as 1 quarter previously. In reality, this is not enough time to implement effective mine planning strategy.

Inadvertently this investigation discovered well performing models at lags of 1, 2, 3 and 9 quarters. This was considered as they had AIC and SIC values below 9. After being stitched together they created The Quilt, demonstrating the 95% prediction interval spread from the best performing models at the lags before mentioned.

The increase in spread is noticed from lags 1 to 3 quarters. There was a noticeable lack of data for lags of 4 to 8 quarters which, arguably is when the ability of operations to implement mine planning strategies are higher. It is recommended that further investigations focus modelling around these equations as they would be more relevant to industry.

This investigation was concerned with the idea of value. Investigating the ability to predict commodity prices is relevant as understanding how they move will enable mining operations to be proactive in implementing strategies to maximise value. This is in contrast to what has been observed in the industry at the moment with mining operations reacting to price changes when they are respectively deemed significant enough to do so.

Understanding how commodity prices will move at certain timeframes into the future will allow more dynamic cut off grades as well as more dynamic reserves. Albeit, there is a mismatch with the time taken to implement strategy as this notion is more concerned with when sales are received for the product although the mine development up until that point likely takes months, or even years.

Further investigations may be interested in the value to be made, or not lost, by having more dynamic cut off grades and mine plans. Unfortunately, the ability to do this is constrained by numerous things. This includes equipment selection, which is a significant cost with units generally being kept for long periods of time, and processing plants, which generally have a specific range of throughput they can handle to maximise recovery. Exploring the value to be maximised by having more flexible and dynamic strategies receptive to changes in commodity prices is an interesting thought well worth investigating but beyond the scope of this investigation.

The data used in this investigation was believed to be suitable with no significant tangible variables considered to be missing.

The price of Iron Ore data collected was quarterly and ranged from September 2009 to December 2018. Ideally this data set would have been larger but the chosen period frequency was not negotiable and only annual data from the source could be retrieved before this date. Rolling monthly data for a quarter could have been used to include more observations and hopefully demonstrate more robust relationships when modelling but the release of sales data is only done quarterly and often delayed at that so not doing this was not considered regrettable.

Total Sales All TSA was consistently a significant independent variable as it was present during some of the better performing models, particularly the Leading Indicator model with two independent variables alongside Crude Steel Production in China CSC. This was despite having quite a short data set, ranging from Dec 2009 to Dec 2018. The length of the data set was constrained by the available quarterly sales information from Rio Tinto; if Rio was excluded the data set could have been longer. It was presumed that it was better to use a short dataset with all the major producers as opposed to a longer data set with less producers. This would be a more robust approach as, if one of the producers suffered a lack supply or even an oversupply at a particular period, it would be better reflected in the data set.

Crude Steel Production in China CSC was also a consistently significant independent variable in modelling and was present for most forms used in this investigation. The data set used in this investigation was of satisfactory length. Admittedly, data retrieved for this variable as well as Crude Steel Production in the World CSW stretched further back than the data used in this investigation. That was because of China's significance noted in the Literature Review conducted. It was therefore decided that to ensure that this significance was not lost on the data set data was taken only from June 2008 to December 2019. This was because before June 2008 China's share of world crude steel production dropped below 38%. Since that time, China's share has increased and sustained at approximately half of world crude steel production.

Generally, both independent variables for Real GDP Growth, for USA and China, were not considered significant during modelling; they often failed hypothesis testing at a significance level of p-value=0.10 and so it could not be determined that their coefficients were not equal to 0. This may not conclusively demonstrate that they are not relevant to the price of Iron Ore but rather that the form of the model and variables was not suitable. The Real GDP Growth data used in this investigation was the percentage change. The Literature Review conducted demonstrated that most scholarly investigations modelled GDP using the natural log to demonstrate growth rates. Using the natural log in this investigation may have increased the significance of variables trying to relate GDP of both USA and China to the price of Iron Ore.

The oil price for both West Texas Intermediate and Brent Crude were used in the hope that they may be indicators. These two variables did return significant correlation values but were not deemed to be decent independent variables considering that they were not included in any of the best performing models. Their significance to the price of Iron Ore may be real but this investigation has not discovered any reason to believe so. This may be investigated further though if more variables are included in model forms. It is likely they are relevant to the price of Iron Ore but need to be modelled with a greater number of independent variables alongside them.

The Copper price was included in this investigation not because it was believed that it has a tangible effect on the price of Iron Ore but rather because it is used frequently in the finance industry as an indicator for the health of an economy. This is the same for the Federal Funds Rate which was believed to be more relevant to the Iron Ore market but not likely to be a direct, immediate relationship. This is partly supported by the optimum lag being 20 quarters, which is a significant time.

These variables, as well as others in this investigation, were not present in any well performing models but that may be more of a reflection on the models that they were included in as opposed to how significant they are to the price of Iron Ore. This investigation purposely used a minimum amount of independent variables when constructing models as it was noticed that there was significant correlation between them. Including independent variables that have strong relationships would alter the relationship that regression would estimate that they have respectively with the price of Iron Ore. This is also an assumption of multiple linear regression modelling that independent variables aren't random but also aren't exact linear functions of other independent variables. In reality though, it may be beneficial to increase the number of independent variables used when constructing models even if there is suspected collinearity as long as caution is used and the risks are understood.

The legitimacy of the raw data collected for this investigation may also be questioned. In particular, the Real GDP Growth Rate of China was particularly difficult to retrieve at quarterly intervals. Several reputable sources did not possess any relevant data and other sources that did possess it demonstrated different values. This may be due to the lack of transparency around the Chinese economy.

It was assumed that data for relevant independent variables could be summed each quarter and be relevant to the price of Iron Ore in a monthly average at the end of later quarters. This reduced the price of Iron Ore observations available considerably but was necessary considering that the scope of this investigation included quarterly data. Increasing the scope of this investigation to consider monthly prices would enable more observations. This could still be done considering quarterly independent variables but rolling their values as opposed to not. Undertaking this action may not be possible for all data retrieved in this investigation but would considerably increase the number of observations used and hopefully reveal more robust relationships. Unfortunately, having more frequent price observations introduces more speculation and may make modelling with tangible independent variables more difficult.

The criterion for model selection were believed to be suitable. Generally, they unanimously pointed to the best models but in the rare occasion that they didn’t the SIC was relied upon to make the ultimate decision. This does not seem unreasonable and does not represent any issues with this investigation.

11. CONCLUSION

In conclusion, this investigation was a success considering that it achieved its aim of identifying key drivers to model the price of Iron Ore and forecast.

This investigation used quarterly price of Iron Ore data ranging from September 2009 to December 2018. The independent variables used for modelling included the price of Iron Ore IOP, Crude Steel Production in China CSC, Crude Steel Production in the World CSW, Real GDP Growth Rate USA GDPU, Real GDP Growth Rate China GDPC, Oil Price West Texas Intermediate OPW, Oil Price Brent Crude OPB, Copper Price CP and Federal Funds Rate FFR. The model forms considered included the Autoregressive AR, Leading Indicator LI, Leading Indicator variation with two independent variables LI2, Autoregressive Distributed Lag ARDL, Autoregressive Moving Average ARMA and the Autoregressive Moving Average with a leading indicator ARMAX.

The best way to model the price of Iron Ore was using a variation of the Autoregressive Moving Average model with the inclusion of a leading indicator. This model used the price of Iron Ore from 1 quarter ago, the error term value from 3 quarters ago, and Crude Steel Production in China from 9 quarters ago as well as a constant. This relationship is demonstrated with the equation,

$$IOP_t = 134.0886 + 0.566401\,IOP_{t-1} + 0.477979\,u_{t-3} - 0.524117\,CSC_{t-9} + e_t.$$

Equation 33

This was believed to be the best model considering it had the highest R^2, lowest Akaike's Information Criterion AIC and the lowest Schwartz Information Criterion SIC with values of 0.922061, 7.937589 and 8.117161 respectively.

It was then assumed that this model would be most suitable for forecasting and the price of Iron Ore in March 2019 was predicted to be US$64.42/t. In reality it was US$85.75 which represents a total error of negative US$21.33/t.

The significant difference in forecast and reality is likely caused by the tailings dam failure in mid-January 2019 of Vale's Brumadinho in Brazil. Production ceased immediately causing a significant reduction in supply to the market with the price of Iron Ore subsequently lifting.

This demonstrates the greatest weakness of this model, and all others included in this investigation, in that they work best in more stable and stationary times and can't anticipate major shocks to the market like the one recently observed.

12. RECOMMENDATIONS

The relevant recommendations for future investigations include to:

- Include a variable that considers the desperation in the market, or lack of; possibly the net of supply and demand of Iron Ore,

- Focus on understanding how to best model the price of Iron Ore at longer lags, from 4 quarters onwards, which would be a more realistic timeframe for operations to implement effective mine planning strategies,

- Include different model forms as well as non-linear relationships for variables in the models,

- Reduce the number of observations used to estimate the relationships within the model to allow for more opportunities to compare forecast models with reality, and

- Investigate the value to be gained, or not lost, by being more receptive to changes in commodity prices and having more flexible cut off grades.

13. REFERENCES

Bazhanov, A, Difficulties in the forecasting of iron ore price: a review [online], Available from < www.mpra.ub.uni-muenchen.de> [Accessed: 19 October 2018].

Chen, J, 2019, Futures [online], Available from <www.investopedia.com> [Accessed: 16 October 2018].

Chen, J, 2019, Spot price [online], Available from <www.investopedia.com> [Accessed: 16 October 2018].

Coates, B, Devlin, W, Woods, S, 2012, Commodity price volatility [online], Available from <treasury.gov.au> [Accessed: 18 October 2018].

Department of Industry, Innovation and Science, 2018, Resources and Energy Quarterly [online], Available from <www.publications.industry.gov.au> [Accessed: 28 September 2018].

Economics Online, 2019, Perfect competition [online], Available from <www.economicsonline.co.uk> [Accessed: 18 October 2018].

Geology.com, 2019, Facts about copper [online], Available from <www.geology.com> [Accessed: 8 April 2019].

Geology.com, 2019, Iron Ore [online], Available from <www.geology.com> [Accessed: 14 October 2018].

Geoscience Australia, 2019 [online], Available from <www.ga.gov.au> [Accessed: 14 October 2018].

Hambur, J, La Cava, G, 2018, Do interest rates affect business investment? Evidence from Australian company-level data [online], Available from <www.rba.gov.au> [Accessed: 30 October 2018].

Investopedia, 2018, How the iron ore market works (Vale, Rio) [online], Available from <www.investopedia.com> [Accessed: 1 April 2019].

Kramer, L, 2019, What is GDP and why is it so important to economists and investors? [online], <www.investopedia.com> [Accessed: 12 October 2018].

Lannin, S, 2017, Iron ore: Small miners under big pressure if iron ore price falls [online], Available from <www.abc.net.au> [Accessed: 10 October 2018].

Market Index, 2019, Iron Ore [online], Available from <www.marketindex.com> [Accessed: 2 April 2019].

McHugh, B, 2016, Iron ore price heavily influenced by trading as well as supply and demand says industry veteran [online], Available from <www.abc.net.au> [Accessed: 8 October 2018].

Oil Price, 2009, Why copper is such an important economic indicator! [online], <www.oilprice.com> [Accessed: 8 April 2019].

World Steel Association, 2019, About Steel [online], Available from <www.worldsteel.org> [Accessed: 1 April 2019].

14. APPENDICES

14.1 RAW DATA

Table 11. Raw Data: IOP, TSA, CSC, CSW, GDPU and GDPC

Date	IOP, US$/t	TSA, Mt	CSC, Mt	CSW, Mt	GDPU, %	GDPC, %
Sep 2005					0.891	
Dec 2005					0.631	
Mar 2006					1.330	
Jun 2006					0.234	
Sep 2006					0.155	
Dec 2006					0.852	
Mar 2007					0.235	
Jun 2007					0.573	
Sep 2007					0.543	
Dec 2007					0.608	
Mar 2008					-0.575	
Jun 2008			138.019	359.001	0.516	
Sep 2008	60.800		127.318	338.309	-0.541	
Dec 2008	69.980		108.391	267.602	-2.164	
Mar 2009	64.070		126.934	264.818	-1.123	
Jun 2009	71.660		139.455	286.582	-0.144	
Sep 2009	80.710		153.836	324.648	0.364	
Dec 2009	105.250	166.300	148.652	331.544	1.098	
Mar 2010	139.770	156.600	155.883	340.965	0.385	
Jun 2010	143.630	163.900	164.587	363.905	0.922	
Sep 2010	140.630	171.600	151.665	340.646	0.737	
Dec 2010	168.530	180.800	151.675	346.213	0.502	
Mar 2011	169.360	154.200	173.594	376.289	-0.240	2.600

Jun 2011	170.880	170.200	179.209	384.339	0.715	2.300
Sep 2011	177.230	186.700	174.752	385.499	-0.028	1.800
Dec 2011	136.450	193.700	156.720	355.959	1.159	1.500
Mar 2012	144.660	162.000	174.197	377.442	0.783	2.000
Jun 2012	134.620	186.400	182.022	388.417	0.430	2.100
Sep 2012	99.470	191.400	178.342	376.005	0.135	1.800
Dec 2012	128.870	207.900	174.223	368.358	0.114	2.000
Mar 2013	139.870	177.600	197.324	395.525	0.886	1.900
Jun 2013	114.810	201.100	202.396	405.890	0.123	1.800
Sep 2013	134.190	220.900	203.038	282.122	0.783	2.100
Dec 2013	135.790	227.800	198.226	400.615	0.798	1.600
Mar 2014	111.830	211.000	202.839	408.044	-0.252	1.800
Jun 2014	92.740	242.800	210.045	420.542	1.254	1.800
Sep 2014	82.270	249.400	206.508	411.291	1.209	1.800
Dec 2014	68.800	264.300	200.490	404.526	0.472	1.700
Mar 2015	56.940	239.400	200.011	401.912	0.823	1.800
Jun 2015	62.290	260.900	207.808	411.937	0.825	1.800
Sep 2015	56.430	274.500	198.900	395.929	0.240	1.700
Dec 2015	39.600	277.800	193.810	386.327	0.100	1.500
Mar 2016	55.520	248.000	191.723	386.652	0.385	1.400
Jun 2016	51.360	267.900	209.394	410.862	0.566	1.900
Sep 2016	56.670	269.500	203.548	401.809	0.478	1.700
Dec 2016	79.430	286.600	202.018	402.966	0.438	1.600
Mar 2017	87.200	248.500	200.762	409.494	0.443	1.500
Jun 2017	57.860	266.900	218.267	427.292	0.740	1.800
Sep 2017	69.690	275.000	220.442	431.859	0.698	1.700
Dec 2017	71.280	286.600	205.560	419.420	0.568	1.600
Mar 2018	69.720	262.000	210.797	426.031	0.550	1.500
Jun 2018	65.110	280.200	238.021	455.187	1.024	1.700

Sep 2018	68.280	280.800	242.412	455.364	0.829	1.600
Dec 2018	67.820	281.500	236.294	452.427	0.641	1.500

Table 12. Raw Data: OPW, OPB, CP and FFR

Date	OPW, US$/b	OPB, US$/b	CP, US$/t	FFR, %
Sep 2003	28.280	27.100	1790.000	1.010
Dec 2003	32.120	29.880	2202.000	0.980
Mar 2004	36.730	33.800	3000.000	1.000
Jun 2004	38.030	35.190	2689.000	1.030
Sep 2004	45.930	43.380	2903.000	1.610
Dec 2004	43.230	39.650	3140.000	2.160
Mar 2005	54.170	53.080	3379.000	2.630
Jun 2005	56.390	54.310	3530.000	3.040
Sep 2005	65.540	62.980	3851.000	3.620
Dec 2005	59.410	56.750	4577.000	4.160
Mar 2006	62.890	62.250	5124.000	4.590
Jun 2006	70.930	68.860	7223.000	4.990
Sep 2006	63.820	62.770	7623.000	5.250
Dec 2006	62.000	62.310	6681.000	5.240
Mar 2007	60.600	62.140	6465.000	5.260
Jun 2007	67.490	71.320	7514.000	5.250
Sep 2007	79.910	77.130	7671.000	4.940
Dec 2007	91.360	91.450	6631.000	4.240
Mar 2008	105.470	103.280	8434.000	2.610
Jun 2008	133.930	133.050	8292.000	2.000
Sep 2008	103.940	99.060	6975.000	1.810
Dec 2008	41.440	41.580	3105.000	0.160

Mar 2009	47.975	46.839	3771.000	0.180
Jun 2009	69.584	68.594	5013.000	0.210
Sep 2009	69.443	67.687	6196.000	0.150
Dec 2009	74.487	74.670	6977.000	0.120
Mar 2010	81.250	79.275	7467.000	0.160
Jun 2010	75.354	74.838	6502.000	0.180
Sep 2010	75.260	77.787	7730.000	0.190
Dec 2010	89.223	91.797	9153.000	0.180
Mar 2011	102.916	114.441	9503.000	0.140
Jun 2011	96.254	113.758	9067.000	0.090
Sep 2011	85.617	110.879	8300.000	0.080
Dec 2011	98.612	107.970	7559.000	0.070
Mar 2012	106.150	124.929	8471.000	0.130
Jun 2012	82.360	95.589	7428.000	0.160
Sep 2012	94.606	113.383	8088.000	0.140
Dec 2012	88.191	109.640	7966.000	0.160
Mar 2013	93.118	109.240	7652.000	0.140
Jun 2013	95.790	103.110	7000.000	0.090
Sep 2013	106.314	111.621	7159.000	0.080
Dec 2013	97.902	110.634	7215.000	0.090
Mar 2014	100.573	107.406	6650.000	0.080
Jun 2014	105.242	111.868	6821.000	0.100
Sep 2014	93.349	97.336	6872.000	0.090
Dec 2014	59.100	62.163	6446.000	0.120
Mar 2015	47.784	55.791	5940.000	0.110
Jun 2015	59.805	62.346	5833.000	0.130
Sep 2015	45.481	47.235	5217.000	0.140
Dec 2015	37.241	37.722	4639.000	0.240

Mar 2016	37.774	39.071	4954.000	0.360
Jun 2016	48.750	48.480	4642.000	0.380
Sep 2016	45.170	46.190	4722.000	0.400
Dec 2016	52.010	54.070	5660.000	0.540
Mar 2017	49.580	51.970	5824.000	0.790
Jun 2017	45.170	46.890	5719.000	1.040
Sep 2017	49.890	55.640	6575.000	1.150
Dec 2017	57.930	64.040	6830.000	1.300
Mar 2018	72.780	66.680	6761.000	1.510
Jun 2018	75.910	67.300	6872.000	1.820
Sep 2018	79.150	70.060	5994.000	1.950
Dec 2018	57.760	49.080	6004.000	2.270

14.2 AUTOREGRESSIVE MODELLING RESULTS TABLES

Table 13. Autoregressive AR Summary Or R-Squared, AIC And SIC With A Constant

Dependent Variable	Lag, Quarter	R^2	AIC	SIC
IOP	1	0.826144	8.528985	8.612574
IOP	2	0.675028	9.168323	9.252767
IOP	3	0.558323	9.482496	9.567807
IOP	4	0.334461	9.907462	9.99365
IOP	5	0.192797	10.12298	10.21005
IOP	6	0.115454	10.24463	10.33261
IOP	7	0.047691	10.32419	10.41307
IOP	8	0.023533	10.34795	10.43774
IOP	9	0.009412	10.36364	10.45434
IOP	10	0.001774	10.30675	10.39836
IOP	11	0.000027	10.22383	10.31635

IOP	12	0.002121	10.10564	10.19905
IOP	13	0.013565	9.901863	9.99616
IOP	14	0.07725	9.788885	9.884042
IOP	15	0.160132	9.596376	9.692364
IOP	16	0.291029	9.350472	9.447248
IOP	17	0.622113	8.750976	8.848486
IOP	18	0.768936	8.184842	8.283013
IOP	19	0.733961	8.140419	8.239157
IOP	20	0.68681	8.243805	8.342991

Table 14. Autoregressive AR Summary Of R-Squared, AIC, And SIC Without A Constant

Dependent Variable	Lag, Quarter	R^2	AIC	SIC
IOP	1	0.818106	8.525406	8.5672
IOP	2	0.646208	9.203293	9.245515
IOP	3	0.500733	9.553776	9.596431
IOP	4	0.17778	10.06624	10.10934
IOP	5	-0.077001	10.35728	10.40082
IOP	6	-0.238712	10.52583	10.56982
IOP	7	-0.420426	10.66687	10.71131
IOP	8	-0.52592	10.73554	10.78043
IOP	9	-0.60694	10.78682	10.83217
IOP	10	-0.687205	10.7691	10.8149
IOP	11	-0.752253	10.72025	10.7665
IOP	12	-0.828102	10.64437	10.69108
IOP	13	-0.959398	10.51919	10.56634
IOP	14	-1.163425	10.56955	10.61712
IOP	15	-1.342216	10.54791	10.5959

IOP	16	-1.489384	10.52952	10.57791
IOP	17	-1.721675	10.64538	10.69414
IOP	18	-1.839776	10.6103	10.65938
IOP	19	-1.960121	10.4628	10.51217
IOP	20	-1.983224	10.40685	10.45644

14.3 ERRORS FROM AUTOREGRESSIVE AR1 MODELLING

Table 15. Error Values From Autoregressive Modelling

Date	AR Error, US$/t
Dec 2008	9.18
Mar 2009	-5.91
Jun 2009	7.59
Sep 2009	9.05
Dec 2009	24.54
Mar 2010	34.52
Jun 2010	3.86
Sep 2010	-3
Dec 2010	27.9
Mar 2011	0.83
Jun 2011	1.52
Sep 2011	6.35
Dec 2011	-40.78
Mar 2012	8.21
Jun 2012	-10.04
Sep 2012	-35.15
Dec 2012	29.4
Mar 2013	11
Jun 2013	-25.06

Sep 2013	19.38
Dec 2013	1.6
Mar 2014	-23.96
Jun 2014	-19.09
Sep 2014	-10.47
Dec 2014	-13.47
Mar 2015	-11.86
Jun 2015	5.35
Sep 2015	-5.86
Dec 2015	-16.83
Mar 2016	15.92
Jun 2016	-4.16
Sep 2016	5.31
Dec 2016	22.76
Mar 2017	7.77
Jun 2017	-29.34
Sep 2017	11.83
Dec 2017	1.59
Mar 2018	-1.56
Jun 2018	-4.61
Sep 2018	3.17
Dec 2018	-0.46